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

Review of the Comparative Survival Study Draft 2014 Annual Report

Richard Alldredge
Kurt Fausch
Alec Maule
Katherine Myers
Robert Naiman
Greg Ruggerone
Laurel Saito
Dennis Scarnecchia
Steve Schroder
Carl Schwarz

ISAB 2014-5 Updated
Original: October 14, 2014
Revised: October 29, 2014 (see blue text on page 5)

ISAB Review of the Comparative Survival Study (CSS) Draft 2014 Annual Report

Contents

ISAB Review of the Comparative Survival Study (CSS) Draft 2014 Annual Report

I. Background

The Northwest Power and Conservation Council's 2009 amendments to the Columbia River Basin Fish and Wildlife Program call for a regular system of independent and timely science reviews of the Fish Passage Center's (FPC) analytical products. These reviews include evaluations of the Comparative Survival Study's draft annual reports. The Independent Scientific Advisory Board's (ISAB) reviews began four years ago with the evaluation of the CSS's draft 2010 Annual Report (ISAB 2010-5), followed by a review of the draft 2011 Annual Report (ISAB 2011-5), the draft 2012 Annual Report (ISAB 2012-7), and most recently a review of the draft 2013 Annual Report (ISAB 2013-4). This ISAB review of the draft 2014 CSS Annual Report is the ISAB's fifth review of CSS annual reports in response to the Council's 2009 Program.

II. Summary

This ISAB evaluation begins by suggesting topics for further CSS review, then provides comments on each chapter of the CSS Draft 2014 Annual Report, and ends with editorial suggestions. Most of the CSS's draft 2014 report is an annual update of information in previous years' reports.

Overall, this year's report was well done. Many of the previous reviews' suggestions have been implemented, making current reports much easier to read and digest. Our review, therefore, focuses on new information presented. For all chapters, except Chapter 2 (Life Cycle Model), the ISAB therefore had, at most, minor concerns. However, for the new material in Chapter 2, the ISAB had significant questions and concerns about the methodology used to combine information from multiple sources, particularly the approach to combine likelihood and Bayesian paradigms. The ISAB suggests that, in the final 2014 annual report and future project documents, the authors provide additional rationale and justification for the approaches used to combine information from multiple sources. Further, results presented in Figures 2.7 onwards produced some unexplained patterns. For example, why did the models tend to consistently underestimate observed survival in freshwater and overestimate survival in the ocean? These patterns may indicate a systematic error in the models. If so, what is the cause of this, and can it be addressed?

The results of Chapter 6 (tag-effects experiment) are very preliminary. However, by the end of 2014, half of the adult production expected from this study will have returned to the Carson Hatchery. Next year's report should also include information on the number of adults produced and the tag loss rates experienced by each tag group. As the SARs become available in future years, this should result in an important peer-reviewed publication.

Overall, the ISAB believes the CSS reports provide valuable information and analyses that are fundamental for evaluating the survival of Columbia Basin salmonids.

III. Suggested Topics for Further Review

(a) **Hypotheses on mechanisms regulating smolt-to-adult return rates (SARs).** The previous ISAB review had several suggestions on mechanisms regulating SARs, to which the CSS response was that it was "out of scope" of their mandate. That being the case, and since the CSS reports provide much of the data necessary for investigation of these questions, the ISAB strongly recommends that the information be available in digital formats. This will allow researchers interested in mechanisms regulating SARs to have easy access to the data. With the time series now approaching 20 years for some stocks, methods such as those in Pyper et al. (2001) may be suitable to investigate correlational patterns in migration timing, in-river survival, ocean survival, and SARs among stocks over large spatial scales to infer common effects on disparate stocks.

(b) **Life-cycle modeling questions and Fish and Wildlife Program SAR objectives**. The previous ISAB review provided two questions that the life-cycle model should be able to answer.

(1) "*What changes in stream productivity* [salmonid productivity in rearing streams] *would be required to achieve population recovery if hydrosystem survival were to remain at the status quo?*"

(2) "*What changes in hydrosystem survival would be required to achieve a 20% increase in population abundance by a particular time in the future?*"

The CSS response agreed that these were important questions and felt that the next iteration of the model would be able to answer at least one of the questions. However, potential problems in the current model's implementation may preclude answering either question at this time. The ISAB believes that it is still important to answer these questions in subsequent years.

(c) **New PIT/CWT study**. The current study design, described in Chapter 6, may only detect the minimal potential adverse PIT tag effects experienced by fish in the Basin. In this experiment, larger fish are tagged before smoltification and fish are allowed to recover for several months from tagging stress. However, many field studies in the Basin apply PIT-tags to migrating smolts (some as small as 60 mm) captured in the field during warm water periods and release them soon afterward. To fully appreciate the potential impacts of PIT tags on survival and to adequately link these effects to fish tagged, a study paralleling those conditions should be performed.

IV. Comments on the draft CSS 2014 Annual Report by Chapter

Chapter 1. Introduction

The introduction to the 2014 CSS Annual Report is a minor revision to those of previous versions. The new Figure 1.1 is very helpful. The ISAB reiterates its previous suggestion for adding a table with a historical timeline of key objectives and results from past CSS work. It would be helpful to summarize these in a timeline with separate columns for changes in transportation, tagging, and so forth. The table could also include citations to past reports and publications containing detailed information on past results, which would be useful to those not familiar with CSS activities.

Specific comments and questions

p.4, l.25. Since PIT-tagging must be done on larger fish, this must have some impact on interpretation. Upgrades in PIT-technology now allow smaller fish to be tagged. Has the tagging program changed the distribution of sizes of fish that are tagged over time? Some commentary here would be useful.

p.10. Bootstrapping. Fish are released as groups and not individually? Does the bootstrapping take into account the potential for over-dispersion by bootstrapping individual release groups (and combining them), or are all fish pooled before bootstrapping occurs?

p.11, l.24. The sockeye salmon tagging location will be changed in a few years to the Sawtooth Hatchery. There is a danger in changing a long-term program’s location without some overlap for a few years between the two locations. The danger is that temporal effects become completely confounded with location effects. When changes are made in locations in long-term data sets, the ISAB recommends that a few years of data with concurrent operations be collected at both locations so the location effects can be disentangled from the temporal effects.

p.12, l.1. Some data must be available on detections of the delayed-migration fish? How are these data used?

Chapter 2. Life cycle modeling approach to estimating in-river and early ocean survival

The objective of this chapter is to develop a multiple-stock model linking freshwater spawning and rearing (FSR) production and survival to mainstem passage survival and ocean survival. The ultimate goal is to use the model to assess important management scenarios to recover spring Chinook salmon. The investigation uses a long time series of SAR data and smolt per spawner data for up to six populations within the Grande Ronde Major Population Group (GRMPG). In the previous report, three different models were evaluated, but in this report the life cycle modeling efforts are concentrated on the LCX model based on the previous analyses. The choice of this model needs to be periodically reviewed in case this decision was premature and based on "one-time-only" features of the data. The changes in this iteration of the report include (1) integration of PIT-tag information and other variables into the life cycle model, (2) separation of survival into distinct mainstem and early-ocean components, (3) inclusion of additional empirical evidence in model fitting procedures, and (4) allowance of variability in freshwater productivity.

Generally, the more appropriate way to combine information from different types of studies is to simply use the product of the likelihoods from the two studies in the optimization process. If the likelihoods are complex, an approximation would be to assume that the estimates from a study follow an approximate multivariate normal distribution with covariance structure based on the sampling variances and covariances (i.e., the square of the standard errors would form the sampling variances). Again, the product of the (approximate) multivariate normal distribution and the likelihood from the life-cycle model would be used in the optimization process. The methods proposed by the CSS group (page 30, equations 2.16 and 2.17) treat the empirical SARs and in-river S as data points that are compared to estimates of the same from life-cycle model. The equations make no use of the estimates' precision or of any sampling correlation among estimates, across years, or within years of the estimates. The proposed approach requires a much more rigorous justification. The equations also seem to add a process error – it appears that there is the addition of a hierarchical Bayesian component (e.g., the yearly in-river S's come from a distribution around a long-term mean with some year-to-year variability). If so, then the whole modeling structure needs to be revised to add this hierarchical component.

Furthermore (page 30), a Bayesian analysis is run after the maximum likelihood estimates (MLEs) are found. But no information is presented on which parameters have priors or on the actual prior for each parameter. Without this information, it is very difficult to interpret the resulting estimates. If a Bayesian approach is intended, then the preferred approach would be to implement it with full specification of the priors, burn-in, thinning, testing for mixing, testing for convergence, and assessing long-term autocorrelation in the posterior values.

The life-cycle model uses multiple sources of data, but more detail about the data used in the model is required. For example, what is PITPH: powerhouse contact rate? How was survival empirically estimated through the first year at sea? (It is described in published papers, but it should be presented here as a means to show assumptions about early versus late marine mortality). What types of juvenile data are used? Four life history types have been described for spring/summer Chinook populations: 1) fry that emigrate from the natal area in March-June (subsequent rearing area unknown), 2) age-0 smolts that emigrate in May-June and enter the ocean as subyearlings, 3) parr that emigrate from the natal river in July-November but overwinter upstream of Lower Granite Dam, and 4) age-1 smolts that emigrate from the natal river in March-June and migrate to sea as yearlings (Copeland and Venditti 2009). In the Pahsimeroi River, age-0 smolts represent approximately 33% of total smolt production from this watershed. Age-0 smolts (and fry) are present in other Snake River watersheds, but their relative abundance is not known. Abundances of fall parr migrants are greater than that of age-1 natal river smolts in Snake River watersheds (Copeland et al. 2014). How does the life history model consider each of the life history types? It is not clear from the text which life stage was tagged in order to estimate in-river and early-marine survival (i.e., age-0 smolts, fall parr, or natal smolt). Presumably, the values were based on fall parr and natal smolts at Lower Granite Dam. Copeland and Venditti (2009) hypothesized that significant gains in abundance and population productivity of Pahsimeroi River Chinook could be achieved via improved SAR of age-0 Chinook smolts, which are abundant yet seem to have a very low SAR. Migration timing of age-0 smolts is later than yearling smolts produced by fall parr migrants and age-1 natal river smolts.

The results presented in Figures 2.7 onwards show unusual patterns that are unexplained. Why did the models tend to consistently underestimate observed survival in freshwater and overestimate survival in the ocean? There appears to be a systematic error in the models – what is the cause of this unexplained response and can this problem be fixed?

This chapter ends with a conclusion about the benefits of transporting fish. Given the concerns identified above, the ISAB believes that only the results on the benefits from transportation from subsequent chapters should be considered in decision making at this time.

Finally, as noted in previous ISAB reviews, the life-cycle model should be used to explore some key questions. For example, given the observed survival at sea, how much improvement is needed in in-river survival to reach a viable population? This question was partially covered in Chapter 4 specific to overall SAR, but not in-river survival. Or, how many smolts per spawner (production from freshwater) are needed to reach a viable population given current SAR? The latter question could help guide habitat restoration and spawning escapement goals given the density dependence observed in other studies.

Specific comments and questions

General. The CSS report is a valuable data source. The brood table for each population should be published in an appendix to the chapter. How was age at return determined back to 1964– the beginning of the brood tables?

p.20, l.25. Combining information from multiple populations is a worthwhile objective. A Bayesian method would be ideal using a hierarchical model vs. the likelihood approach currently employed in the life-cycle model. The hierarchical model would allow information about parameters from stocks with much data to inform the parameter values for stocks with less data.

p.21, l.31. Why are only years with complete data used for all populations – as far as the ISAB can determine there are no technical challenges in using data from years when only subsets of the populations are measured.

p.21, l.37. What is the difference between the NPH and the PITPH variables, and how they are constructed? More details are needed.

p.22, l.3. If the PITPH and WTT generally track each other, why are both used? Does this co-linearity cause any problems in model fitting and interpretation?

p.22, l.20. This description seems to imply that a Bayesian Hierarchical model is being fit, but the rest of the document does not describe fitting a hierarchical model.

p.22, l.29. A B-H relationship is used only at the smolt-production stage. Timing of migration is not considered, so how will any density dependence of co-migrating stocks be handled?

p.24, l.29. Why is the LCH-PE called a “process-error” model? Process-error has a very specific statistical meaning that does not appear to be relevant here. “Errors in predictions” is also used in an odd fashion. In analogy with regression, it does not matter what the cause of the discrepancy is between the observed and predicted data – the fitting procedure assumes that this random variation can be described and should generally be “minimized” by the fit.

p.24, l.33. There may be serious problems in assuming that the predicted number of spawners is the actual number of spawners used in future model predictions. This is analogous to using the expected value of the regression line as the actual value of the response. This creates an “error-in-variables” problem, and the model fitting is likely inappropriate for this type of model. This will also fail to produce the full variability in future spawners and returns since the model gives a result more close to an average. For this reason, a fully Bayesian approach may be more suitable for forward predictions.

p.25, Figure 2.2. Analyses by Zabel and Cooney (2013) show that these populations exhibit strong density dependence across the life cycle. Given these relationships, the CSS might consider correlating residuals from each recruitment curve against each population and environmental variable rather than Log R/S, as shown in Fig. 2.2. Use of residuals would remove density dependence associated with parent spawners in each population and might provide more representative relationships.

p.26, Equation 2.1. It is not clear which life history types were included in the model. For example, do smolts (M) in equation 2.1 refer to all smolts or only those migrating this year? Relative production of the life history types may be density dependent: more young migrants were produced from greater numbers of spawners according to Copeland et al. (2014).

p.26, Equation 2.3. Where do the survival rates in equation 2.3 come from? Are they the empirical survival rates? Also, how was the proportion of fish transported versus fish in river determined? The report indicates these are estimated from PIT-tag data, but some pointers to the tables where the data are given would be helpful.

p.26, l.3. The LCH-DEV model is actually a process-error model employing the traditional usage of process-error in modeling (i.e., a common year effect among populations).

p.26, l.13. Do the LCH-PE and LCH-DEV models assume that the variance of the productivity parameter over time is zero? If so, why is this done? Equation [2.1] and following have no variability in the actual numbers produced. In the typical Beverton-Holt and Ricker models, there is variability above and below the underlying line describing the relationship. Equation [2.3] does not allow for any stochastic variable. There is no stochastic variability in survival or maturation either. This suggests a preference for a state-space model using Bayesian methods rather than a likelihood formulation.

p.28, Equation 2.9. How was harvest rate calculated for the model? Did it come from CWT data?

p.30, Equation 2.16. This equation can be clarified by improving the notation. The in-river survival is a parameter and the empirical S is data. The ^-indicates that the LCM term is an estimate, but it is a parameter. The empirical S should have the circumflex while the parameter should have none. (The equations are symmetric, so it works out properly in the end). Furthermore, Equation 2.17 uses the SAR, but the likelihood equations on page 26-28 do not include an explicit SAR term. The ISAB is concerned about the way in which this additional information (both the empirical survival and SARs) are being merged with the LCM, and these proposed methods require a rigorous justification.

p.30, l.19. The MCMC performed after the maximum likelihood estimation is an attempt at a Bayesian implementation, but what priors were used for each parameter? Are these appropriate priors?

p.31, Table 2.3. Are the smolt productivity estimates in units of smolts per spawner? If so, how do these values compare with other published data for the region (e.g., Walters et al. 2013)?

p.31, l.13. Presumably, predicted survival values are from models that excluded the environmental variables.

p.33, Table 2.5. MLE and SD (from a Bayesian analysis) are reported. Paradigms are mixed in this table. When MLEs are reported, they should be accompanied with a SE. If a summary of the posterior is being reported, then an SD is appropriate. But, what priors are being used and are these appropriate – particularly for the variance parameters?

p.33, Table 2.6. Bayesian and likelihood (AIC) paradigms are again mixed. The DEV model does not have an extra 43 parameters because there are random variables. The MLE would integrate over these variables (much like random effects in mixed models). A Bayesian model might compute the effective number of parameters using DIC and related methods, but this was not done here either. It is not clear why these terms were counted as parameters.

p.34, Figure 2.5. Estimates and parameters of models are mixed. Data on in-river survival from the PIT-tagging does not depend on the model. Functions of parameters are being plotted.

Figure 2.10. These are more like best linear unbiased predictions (BLUPs) rather than predicted values.

p.39, l.39. Here, the mean of the posteriors are plotted; previously the mode of the posterior was used. Why the switch between these two measures of the posterior distribution?

Figure 2.16. Some of the posterior plots look odd, especially plots (b) and (q). Is there some bounding of parameters in the MCMC runs? Please explain.

Chapter 3. Effects of the in-river environment on juvenile travel time, instantaneous mortality rates and survival

This chapter is similar to the previous year's report. Regression models are used to investigate the influence of environmental and operational covariates on the estimates of instantaneous mortality and fish travel time. A statistics-on-statistics approach is used rather than embedding the covariates directly into the CJS mark-recapture models as suggested in the ISAB's previous reviews. The AIC paradigm (model averaging and so forth) is used to investigate the support for different models relating survival and other variables.

Overall, the text of this chapter is clearly written. Figure 3.1 is a helpful addition to keep track of the relative location of dams. It would be useful to provide a summary of the fish used and not yet used in modeling by identifying attributes such as hatchery or wild, species, and life-history type. A table might be an effective method to present such information.

In regards to the environmental variables on page 58, what is the rationale for using the 7-day window around the median passage date for these variables? Is it an arbitrary choice, or is it related to fish travel times or variability of environmental conditions? It is not clear that "average spill percentage" is an adequate metric to use as one of the environmental variables Justification for using the average value is needed.

If the coefficients of variation presented in Table 3.3 are considered valid measures of model fit, then some discussion should be included of why the models for FTT fit so much better than those for Z and survival. Also, it may be worthwhile to present and interpret the low coefficient of determination values for yearling Chinook compared to other species in Table 3.3.

The results of this chapter are used to support a change in spill regimes as a future experiment. The ISAB agrees that the models presented in this analysis provide a basis for conducting adaptive management experiments as stated on page 70. Furthermore, the CSS provides additional suggestions for future investigations such as those included at the end of Chapter 3. However, it is not clear what specific data or logic suggests that survival would have been greater if spill levels at Ice Harbor had not been reduced during 2005-2013 (p.65, l.17). This assertion needs to be supported with more explanation and justification.

Specific comments and questions

p.55, l.16. The mean FTT is computed only for fish detected at subsequent dams. This assumes that fish not detected are a random sample of fish passing the detector, so that detection does not depend on the speed of the fish or that faster fish do not choose a different route. Is there justification for this assumption?

p.56, l.19. The coefficient of variation for hatchery and wild sockeye is reported as being 41%, but the value in Table 3.1 is 39%. Should these values be the same?

p.58, l.2. Evidence should be provided to support the assumption that –log(S) is log-normally distributed.

p.61, l.1. The coefficient of determination is not a valid measure of the proportion of variance explained for models that lack an intercept term, and hence for many non-linear models. It is not clear that the use of the squared Pearson correlation coefficient is appropriate. Please clarify.

p.66. Several cohorts in the same year are treated as independent, but this assumption may not be true. Consequently, a mixed-effects model is used. R^2 is not well defined for mixed-effect models as the REML approach to model fitting does not provide the usual sums-of-squares used in the computation of R^2. How are these values determined?

Chapter 4. Patterns in Annual Overall SARs

This Chapter updates the CSS time series of smolt-to-adult return rate (SAR) estimates of previous reports with an additional year of data. The same methods are used as in previous years, and so these are not extensively reviewed this year. A new section "SARs vs. Population Productivity" has been added.

Productivity has been defined here as adult returns/spawner. SAR is defined as adult returns/smolt. The number of adult recruits appears on the numerator in both quantities, and so it is not surprising that there is a strong correlation between the two quantities. It is unclear to the ISAB if valid conclusions can be drawn from what may be a mathematical artifact of the definitions.

The authors recognize the need to incorporate density dependence into the examination of population productivity versus SAR (p. 100). They should also examine the level of smolts per spawner needed to achieve viability given current SAR values while also considering the observed density dependent relationship between smolts per spawner and spawners.

Long time series of SARs are available for many stocks. The ISAB suggests that these be used to examine patterns of correlation over time across stocks in order to separate which stocks are experiencing similar environmental effects because of similar feeding, migration, and ocean habitats (see the analyses of Pyper et al. 2001).

Specific comments and questions

p.74. The computation of the SARs is quite complex, using weighting factors and other terms prior to 2006. Exactly what is bootstrapped in the computations? The bootstrapping methods for data after 2006 are much easier to compute because of the change in the way the data are collected. In both cases, is bootstrapping also applied to the estimated smolts produced? More details are needed on the procedures used.

p.75. The SARs are based on estimated adult returns / estimated smolt numbers. Are both the numerator and denominator bootstrapped to get the full uncertainty in the ratio?

p.78, l.16-17. Population productivity is defined as ln((adult recruits to spawning grounds)/(adult spawners)). These are not the same units as presented in Chapter 2. Consistency in units is needed across chapters.

Chapter 5. Estimation of SARs, TIRs, and D for Snake River Subyearling Fall Chinook

The CSS team continues to make progress in including fall Chinook in a manner similar to spring/summer Chinook and steelhead. Specifically, they continue to make progress in considering the possible large portion of late season migrants and high proportion of holdovers that may introduce bias in estimating SARs, particularly for the C_0 group. Attempts to use predicted holdover probabilities to remove individual fish from their cohorts have not been successful, so some release groups are not used in SAR estimation. Simulation to estimate the effect of bias due to holdover fish is an appropriate continuing effort.

More discussion of the low in-river survival of wild Snake River fish is needed (Table 5.9).

The ISAB appreciates the CSS responsiveness to the request to include new groups outside the Snake River for comparison with those released above Lower Granite Dam and the reporting of SARs to Bonneville Dam in comparison to SARs to Lower Granite Dam.

In Table 5.10 the percent increase in SARs with jacks compared to SARs without jacks appears fairly consistent. Perhaps this should be noted in the discussion of the tabular results. It is also interesting that wild fall Chinook (Table 5.9) had a half to a third of the in-river survival of the hatchery groups (Tables 5.7 – 5.9), but ended up with very good SARs (relatively speaking; Table 5.16). Later the authors do discuss the negative correlation between TIR and survival between LGR – BON, but there is only one year of data for the wild fish (Fig 5.5), so the conclusions must be viewed as very tentative

The section describing estimates of TIR does not provide a description of the tabulated estimates of D. Some explanation for why D is not included in the description is needed.

It is stated on page 184 that there appears to be no overall benefit to transportation. Could it also be concluded that there appears to be no overall benefit to not transporting fish?

Specific comments and questions

p.142. The authors claim that predictions of holdover for individual fish was not effective. Why not fit a model similar to zero-inflation models where two latent groups are postulated but all of the data combined are used?

p.161, l.7-8. "Patterns in overall SARs were similar for all groups, with the highest return year being 2008 and the lowest return year 2007 (Figure 5.2)." This statement is not accurate. In some locations, SARs in some years are higher or the same as 2008 and only 3 of 8 graphs have data for 2007.

p.177. Tables 5.34 – 5.38 include "D," but there is no mention of it or its significance in the text.

Chapter 6. PIT-tag and coded-wire-tag effects on smolt-to-adult return rates for Carson National Fish Hatchery spring Chinook salmon

This is a new section in the CSS report, documenting preliminary results from a well-designed experiment but, as noted below, may provide information only about best-case results. The results are interesting, but with only one year of data, are very tentative. As the SARs become available in future years this should result in a good peer-reviewed publication.

The objectives of the study are well-presented. The study also has the ability to examine the effects of tag type on: (a) adult migration timing to the hatchery, (b) adult size when age and sex are the same, and (c) whether trends in tag loss over time are occurring within a tag type. For example, is there any evidence that older fish have higher rates of tag loss than younger maturing individuals, as might be expected if tag loss continued over time? These evaluations would complement those made by Knudsen et al. (2009) on tagged spring Chinook released into the Yakima River. At the end of the 2014 spawning season, half of all the adults released from the project will have returned (3's, 4's and 5's from the 2011 release, 3's and 4's from the 2012 release and 3's from the 2013 release). If efforts to collect such data were not made or cannot be retrieved from hatchery records, we recommend that they be collected on future returning adults.

Like the study performed by Knudsen et al. (2009), this experiment probably represents the best-case scenario for potential PIT tag effects. As Knudsen et al. (2009) mention, they tagged fish in November when water temperatures were cool and the fish were large. Tagged fish were also allowed to recover for 70 or more days prior to being released. The current study takes place under similar circumstances. Relatively large fish (~13 g) were tagged in mid-November and were not released until mid-April. Consequently, the fish were allowed to recover from tagging stress for multiple months, and they were also tagged before they underwent smoltification. Many field studies using PIT tags use migrating smolts (some as small as 60 mm) captured in the field during warm-water periods and release them soon thereafter (Knudsen et al. 2009). To fully appreciate the potential impacts of PIT tags on survival, a study paralleling those conditions should be performed.

Results of the current study indicate that dual-tagged fish survived at comparable rates to fish with a single tag (either PIT or CWT) in the hatchery and from release to Bonneville Dam (i.e., over relatively brief periods of time). Yet from Bonneville Dam to return, dual-tagged fish survived at 1/3rd the rate as the other two types of tagged fish. Although the results are of a preliminary nature, some discussion on why this might be happening should be included in the report.

Specific comments and questions

p.188, l.13. The authors claim that ANCOVA will be used, but the subsequent analysis on the number of returns from a given number of releases will be a variant of Poisson or logistic model. More details are needed.

p.190, l.12. No evidence of a difference was found, but how was this hypothesis tested? Was a likelihood ratio test used with a paired-release design?

p.185, l. 27-28. It is unclear whether PIT-, CWT- and PIT&CWT fish are being reared together (i.e., in the same raceways or separate, in segregated raceways). If they are being reared together, are the same proportions of each type of fish present in all the raceways? This should be clarified because of possible performance effects due to rearing location and release dates.

p.185, l.39-41. The results of a power analysis are provided. However, it was unclear why such an analysis would indicate that one type of tag (CWT) would need 75,000 individuals and the other two types, PIT only and PIT-CWT should have 15,000 representatives per year. Is this differential related to future detection rates?

p.186, l. 9-10. Again, it is unclear whether the fish from each tag group are being reared together or separately.

p.186, l.10. Additional information about how the releases were made would be useful. For example, were they forced or volitional? Were they made during daylight or darkness? Over how many days did the releases occur? If the releases occurred over multiple days were equal proportions of fish from each tag group released on each day?

p.186, l. 16-17. Some explanation is needed on why it was necessary to euthanize dual-tagged fish. Was it because the detection devices used were unable to distinguish between PIT and CWT tags, making false positives likely?

p.187, l. 5-16. Given the difficulties associated with the 2011 release (e.g., not keeping track of tagged fish mortality, tag shedding during the rearing period and poor PIT tag detection efficiency at release), has any thought been given to adding one more year to the study so that more accurate assessments of release numbers/tag group and tag loss during the hatchery rearing period would be available? An additional year would meet the requirements of their power analysis. A description is needed of how fish count information obtained when the fish were adipose clipped and how subsequent mortality and tag shedding data were used to estimate the numbers of fish in each tag group at release. Also, were any comparisons made between the estimates generated by keeping track of rearing mortalities and tag shedding verses those derived from the PIT tag array established in the release channel? If the two types of estimates are close, it would provide some confidence that the estimates of CWT fish leaving the hatchery are reliable.

p.187, l.15-16. As mentioned above, more information is needed about how the staggered releases were made and whether each tag group was proportionately represented at each release.

p.197, l. 17-21. Juvenile survival estimates from tagging to release in the hatchery and from release to Bonneville were determined by two different methods. As currently written, it appears as if just one method is being used. The text should be clarified here.

p.188, l. 1-4. At the time the current CSS report was written, about a third of the adults expected from the project had returned to the Carson Hatchery. Consequently, some assessments of tag loss must have been made on these fish. Yet none of these data is presented in the report. It would be useful if a table could be produced that indicates how many fish have returned (by broodyear) to the hatchery and what the tag retention rates have been for the three tag groups. That information would help answer a number of questions (e.g., how many adults were used to generate the preliminary SARs presented and did females retain their PIT tag as they became mature or were some lost after they arrived at the hatchery). Also, a comparison of tag retention rates at the time of release and at maturation could be made and discussed. As well, it would be possible to see if tag loss varies by age or sex, particularly for females.

p.192. Some further discussion or speculation is needed on why the investigators think the SAR values of the dual marked group are so low relative to the other groups.

Appendix A: (SR), SAR, TIR, and D for Snake River Hatchery and Wild Spring/Summer Chinook Salmon, Steelhead, and Sockeye

This Appendix presents the methodology for the computation of the SR, SAR, TIR, and D parameters along with extensive tables of results. The Appendix is an update of similar material from previous years and so was not reviewed in depth. New diagrams showing the relationship of the statistics proved quite helpful in understanding the various equations.

p.A.15, l.28. Was any adjustment made for multiple testing? Perhaps the 90% CIs need to be adjusted to account for multiple hypothesis tests to reduce the number of false positives (i.e., a large number of statistical tests are done by examining if the 90% CI for D does not overlap the value of 1 indicating that differential delayed mortality effects of transportation were detected). On average, without a correction for multiple testing, about 10% of these tests will find evidence of delayed differential mortality even if it does not exist.

p.A.51. Figure A.18. How was the aggregated hatchery rate computed?

V. Editorial Comments

Glossary. NPH and PITPH seem similar, and it is not clear how these differ.

Glossary. Why two different abbreviations for sockeye? There are also several other terms with two different abbreviations. Please unify into a single abbreviation. Additionally, two different abbreviations are used for Little Goose Dam.

Glossary. Survival rate is defined in terms of (possibly different) time intervals. Either standardize a rate to a per year basis or instantaneous mortality basis (Z which is standardized to a per day basis) or change to "survival probability" rather than "rate."

p.2, l.8. "The number … detected … declines over time." Not necessarily so because of stochastic variability. Add "… declines, on average, over time."

Figure 1.2. Use different symbols, in addition to color to differentiate between different dam types, to help color-blind readers or those using black-and-white printers to differentiate the map. Please review all figures that use color to avoid aforementioned problems.

p.6, l.43. "This method is used to estimate a population of PIT-tagged smolts surviving to the tailrace of Lower Granite Dam and their subsequent survival through the hydrosystem." This sentence reads awkwardly.

p.7, l.30. “researchers submitted groups of PIT-tagged Snake River fish.” Awkward wording. How do the actions before and after 2006 differ? In both cases, are codes used to select fish?

p.7, l.27. This paragraph would benefit from reorganization. There are several separate ideas discussed together here. Tag codes, transportation (“used to begin,” but what happens now?) New stations added. Perhaps a chronology in table format would be useful to summarize changes over time in a linear fashion with separate columns for tagging operations, PIT detectors, transportation and so forth. On page 8, line 30 additional details about the transportation program are provided – all of the information about changes in transportation should be compiled together.

p.7, l.44. Consider adopting a consistent convention for place names and locations for adult detectors. For example, LGR and LGR-AD or BON and BON-AD for the dam and adult detectors on the dam respectively, rather than different looking abbreviations.

p.9. A diagram would be helpful describing the different groups of fish and where TIR and D occur. This suggestion was made in previous years as well.

p.14, Table 1.1. Here and all other tables would be easier to read if numbers were right-justified in columns and character data (e.g., species) were left-justified, rather than centered justified, as is the case now. What is the rationale for the row ordering in these tables? Distance from BON? Grouping by ESU?

p.16, Figure 1.4. Adjust the Y-axis from 0 to 100%.

p.19, l.36. Inclusion of empirical evidence is not in the model fitting procedures. It occurs in the likelihood development. The fitting procedures are the same regardless of how the likelihood is constructed. p.20, l.1 also needs to be reworded. These are additions to the likelihood and not the fitting procedures.

p.19, l.37. The powerhouse index and WTT are added to help what features of the model?

p.20, l.17. “… robust …” Do you mean that the small sample sizes give estimates that are highly variable? This would a more accurate wording than robust.

p.24, l.14 “main stem AND ocean” (the AND was misspelled).

p.24, l.20. Is LCH a superset of the LCX model? Not clear from the description.

p.24, l.24. Again use a more descriptive notation. How about LCH-PE and LCH-OE to remind the reader that these are variants of the same model?

p.24, l.33. “The implementation assumes that the predictions are accurate given the rates used." It is unclear what this means.

p.26, The CSS defines 1-salt, 2-salt, 3-salt etc. in the glossary but then uses 1-ocean, 2-ocean, etc. when describing the model. A consistent notation should be used.

p.56, l.14. "We calculated Chi-square adjusted variances (using the $\hat{\mathbf{c}}$ variance inflation factor)." What is meant by a chi-square adjusted variance? Do you mean you computed c-hat using the ratio of the deviance to its degrees of freedom?

p.57, equation 3.3. Z-hat is a random variable, but S and FTT are not and so have no variances. All of the S and FTT in the equation should read as S-hat and FTT-hat.

p.58, equation 3.4. Ditto as for equation 3.3.

p.59, equation 3.6. Reviewers were somewhat surprised that sqrt(Z-hat) was selected; reviewers would have expected a log-transform as well. The log-transform and sqrt() transform have similar form, so reviewers would do the modeling using log(Z).

p.95, Figure 4.14. What are units? In the text it is %, but these appear to be proportions. Why are confidence limits for proportions for some years greater than 1?

p.99, Figure 4.17. The units of the graphs are proportions but the text reports %. Consistency is needed here.

p.100, l. 20. "…abundance (ln (S/S) < 0)…" Some notational revision is needed because S/S = 1. The same notation is used in Figure 4.18. Do these refer to the definition of productivity as defined on page 78?

p.147. Report bias as percentage points or a decimal fraction to avoid confusion between % bias and SAR % (i.e. is a 2% bias in a SAR of 4%, 2% of 4% (i.e., .0008) or .02 on a base rate of .04?).

p.162, l.16. Check use of "apposed."

p.181, l. 8. Check use of "concision."

p.185, l.41. Power analyses require specification of the effect size. What PIT-tag effect can be detected with the current size of the study?

p.185, l.26-27. Change principle to principal.

p.188. Please define "NA" in Tables 6.1 & 6.2.

p.188. l.23-24. These lines indicate that the total numbers of fish released for each group are shown on Table 6.1 (page 189) The current column label for those numbers in Table 6.1 (page 189) is misleading. It appears to indicate the number of fish tagged in each group, not how many were eventually released from the hatchery. It should be changed to something like "Estimated numbers released," or simply "Numbers of fish released." Another column could be added that indicates the number that were tagged and placed into each group in mid-November.

p.188, l.27-29. A table with the data shown in Fig. 6.2 might be a better way to present this information.

p.191 Figure 6.4. Color for coding categories is hard to read by color-blind readers, and it does not photocopy well.

p.A.51. Figure A.18 top panel. It may be best not to use color only to separate the lines into hatchery and wild stocks.

p.A.53. Table A.38. Column labeled "Delta i" should be "Delta AICc."

p.A.53, Figure A.19. The blue color to separate from black and red dots does not photocopy well.

p.A.54. Figure A.19. "The two red data points…" There appears to be only one red data point.

VI. References

Copeland, T., D.A. Venditti, and B.R. Barnett. 2014. The importance of juvenile migration tactics to adult recruitment in stream-type Chinook salmon populations. Transactions of the American Fisheries Society. In press.

Copeland, T. and Venditti, D.A. 2009. Contribution of three life history types to smolt production in a Chinook salmon (Oncorhynchus tshawytscha) population. Canadian Journal of Fisheries and Aquatic Sciences 66: 1658–1665.

Knudsen, C. M., M. V. Johnston, S. L. Schroder, W. J. Bosch, D. E. Fast, and C. R. Strom. 2009. Effects of Passive Integrated Transponder Tags on smolt-to-adult recruit survival, growth, and behavior of hatchery spring Chinook salmon. North American Journal of Fisheries Management 29:658–669.

Pyper, B.J., F.J. Mueter, R.M. Peterman, D.J. Blackbourn, and C.C. Wood. 2001. Spatial covariation in survival rates of Northeast Pacific pink salmon (*Oncorhynchus gorbuscha*). Canadian Journal of Fisheries and Aquatic Sciences 58:1501-1515.

Walters, A.W., T. Copeland, and D.A. Venditti. 2013. The density dilemma: limitations on juvenile production in threatened salmon populations. Ecology of Freshwater Fish 22:508-519.

Zabel, R. and T. Cooney. 2013. Recruits-per-spawner in base years versus current time period—do they differ? Appendix C of the 2014 FCRPS Supplemental Biological Opinion, NOAA Fisheries, January 17, 2014.