FISH PASSAGE CENTER

847 NE 19th Avenue, #250, Portland, OR 97232
Phone: (503) 833-3900 Fax: (503) 232-1259
http://www.fpc.org/
e-mail us at fpcstaff@fpc.org

MEMORANDUM

To: Rich Alldredge, Chair, ISAB
Bill Bradbury, Chair, NPCC
Paul Lumley, Executive Director, CRITFC
John Stein, Science Director, NOAA Fisheries Science Center
Randy Fisher, PSMFC

Michele DeHart

From: Michele DeHart

Date: November 30, 2014

Re: Response to ISAB comments on the Draft 2014 Comparative Survival Study Annual Report

Attached, please find the Comparative Survival Study (CSS) Oversight Committee responses to ISAB comments on the draft 2013 Comparative Survival Study Annual Report. As in past years the ISAB comments are insightful and have improved the report overall. The original comments are presented in italic font followed by the responses in standard font.

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
October 14, 2014

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.

CSS Response: CSS annually updates data used in development of the annual reports. Those data are currently available via the FPC website at the following location: http://www.fpc.org/survival/smolttoadult_queries.html.
If other information is required researchers can contact CSS or FPC to get more detailed data.

(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.*

CSS Response: We are in the process of developing a timeline—it will be necessary to add it to next year's report—possibly as an appendix.

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.

CSS Response: In our version of Chapter 1 p.4 is a figure with no text. However, in response to your comment, tag size for PIT tags has not changed demonstrably in the past 20 years. The CSS is a cooperator on many tagging efforts. Generally, the minimum size at marking for wild fish has been 60mm, as recommended by PTAGIS, and that size has not changed. Nor would changing it greatly affect the current marking that occurs on wild fish.

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?

CSS Response: All fish are pooled before bootstrapping. As far as we understand, a stratified bootstrap procedure, as suggested, would only be appropriate if fish had unequal probabilities of being detected at or surviving to downstream dams or surviving to adult return. We assume this is not the case with groups that are pooled in CSS. Even so, we did run a quick simulation to compare a stratified bootstrap procedure to the normal procedure that assumes independent random sampling for a binomial distributed random variable (x (# Adults) ~ Binomial(N (# Juveniles), p (SAR)) from 3 sub-populations with unequal sample sizes and SAR rates. Coverage determined by the 90% confidence interval for the two procedures were nearly identical.

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.

CSS Response: We agree that the two hatchery data sets should not be combined. Sockeye production is being phased out at Sawtooth Hatchery by smolt migration year 2016, with production (and the CSS mark group) being shifted to newly constructed Springfield Hatchery. In the single year that sockeye will be released from both hatcheries (2015 smolt migration year), both the Sawtooth and Springfield groups will be PIT-tagged.

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

CSS Response: Data are available on delayed migrating juvenile spring/summer Chinook and steelhead. Typically, fish from a single trap marked within one calendar year can migrate out of two different years. These data are used to identify length and date combinations that can be used to identify annual cohorts from each mark location (trap, weir). Based on length and date tagging, fish are removed from analysis if any fish from that time period and size range were detected migrating in the hydrosystem outside the year of interest. The description in the text for steelhead is similar to the process used for wild spring/summer Chinook.

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

CSS Response: In general we have revised Chapter 2 to reflect the responses provided below with the ISAB comments. In general, we concur that the methods required additional clarification, but we also note that our implementation of process and observation errors is more consistent with accepted definitions than may have been clear on reviewer's first reading. We have updated the terminology for clarity.

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.

CSS Response: The likelihoods in equations 2.13–2.17 are combined by taking logarithms of each term and summing the negative log-likelihoods. Minimizing this is equivalent to maximizing the product of the likelihoods, but it is a more stable approach because likelihoods are very small numbers and the product can be so small that the optimization can become insensitive to adjustments in parameter values. As stated on p30 line 17, the sum of the negative logarithms is minimized. Each likelihood term assumes a separate variance term, so in fact the implementation results in a multivariate normal distribution (in log space). Further clarification is provided in text.

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.

CSS Response: Equations 2.13–2.15 treat the variance as a nuisance parameter. The maximum likelihood estimate of the variance is substituted into each equation, effectively integrating the likelihood over the variance. In the case of the in-river survival likelihood (Eq. 2.16), an estimated variance for each year is included, which weights each year inversely with its variance. Further clarification is provided in revised methods.

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.

CSS Response: Clarification of MCMC simulation procedures provided in revised methods.

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? At least three life history types have been described for these populations: 1) fry that emigrate in spring and enter the ocean as subyearlings, 2) parr that emigrate from the natal river in fall but overwinter upstream of Lower Granite Dam, and 3) smolts that emigrate from the natal rivers in spring and migrate to sea. The first two life history types are considerably more abundant at the time of migration than smolts, according to Copeland et al. (2014). How does the life history model consider each of the life history types? Similarly, it is not clear from the text which life stage was tagged in order to estimate in-river and early-marine survival (i.e., fry, fall parr, or natal smolt). Presumably, the values were based on fall parr and natal smolts at Lower Granite Dam. Copeland and Venditti (2009) and Copeland et al. (2014) hypothesized that significant gains in abundance and population productivity could be achieved via improved SAR of age-0 spring Chinook smolts, which are abundant yet seem to have a very low SAR. Migration timing of age-0 smolts is later than yearling smolts.

CSS Response: Clarification of PITPH and empirical survival estimates is included in revised methods. The life history represented by the model is shown in Figure 2.3 and described in Equations 2.1–2.12. In order to be as parsimonious as possible, the modeled life history in this analysis does not include a distinction for upstream versus downstream overwinter survival, and although both parr and smolts were tagged, the in-river survival begins with smolt detections at Lower Granite Dam.

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?

CSS Response: The empirical data for in-river survival are for the Snake River aggregate, but the prediction is for the Grande Ronde / Imnaha aggregate. The likelihood attempted to fit the survival, but since the overall fit needed to also fit the adult returns, only the general pattern of in-river survival variation is visible in Figures 2.7–2.9. The ocean survivals of in-river migrants and transported fish are presented in Figures 2.7–2.9, but they are not included in any likelihood. They are only shown for comparison.

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?

CSS Response: The methods used to obtain these data are described on the reference data repository. The data are the best available estimates of the number of returns of each age to each spawning area.

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.

CSS Response: We agree that a state-space approach would be ideal, but the 2014 life-cycle model development represents an incremental improvement toward that end.

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.

CSS Response: It's true that there is no statistical reason why all populations need to span precisely the same data range, but there were prohibitive problems to overcome in the coding of life cycle models, and the exception handling required to treat each population differently. It's possible, but it would add little to the analysis since the more recent years are more representative of conditions the life cycle model is being developed to help analyze.

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

CSS Response: Clarification of PITPH and empirical survival estimates are included in the revised methods.

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?

CSS Response: There is some correlation, but if there were no benefit in adding WTT to the analysis, the AIC scores would reflect this. Also, if one of the variables were redundant, the coefficient would be estimated to be zero.

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.

CSS Response: It is not strictly a Bayesian Hierarchical model because there is no prior place on the estimation of the deviates. The deviates are merely estimated as annual pulses and imbedded as an autoregressive error term (see Equation 2.2 and note that the change (minor) to how the deviate is implemented as a log-deviate).

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?

CSS Response: Density dependence occurs before the stocks migrate, so timing is irrelevant to the B-H relationship of each population.

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.

CSS Response: Clarification of process error and observation error terminology are provided in the revised methods.

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.

CSS Response: An observation error model is an "error-in-variables" model. The process is deterministic, so the deviates in the fitting represent the observation errors.

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.

CSS Response: Relating the productivities in FSR to environmental variables in similar fashion to Equations 2.10–2.12 is the intended direction of the CSS life cycle modeling.

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).

CSS Response: Smolts at year t+2 migrate past LGR and are the progeny of spawners in year t.

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.

CSS Response: Survival rates in Equation 2.3 come from Equations 2.10-2.12. Table 2.1 contains a complete list of state variables, estimated parameters and demographic rates.

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).

CSS Response: The LCH-DEV model is a process error model with autocorrelation and an estimated temporal deviate. Further clarification provided in revised methods.

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.

CSS Response: The deviates estimated in Equation 2.2 sum to zero in the DEV model only, and infer inter-annual variability in each FSR productivity with an implied mean of zero. Including a likelihood term that minimizes deviates with an estimated variance would make this a state-space model, but this was not implemented for technical reasons that have not yet been resolved. The deviates are set to zero and not estimated in the PE and OE implementations. Since these are not simulation models, variability is not added explicitly as the comment seems to suggest. Variability in the constant annual rates is inferred from the Bayesian MCMC implementation, but that variability is not explicitly an annual deviate. It's an implied uncertainty in the estimated value of the parameter given the data. Implementing a state-space approach to the life-cycle model remains a priority.

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

CSS Response: Harvest rates were calculated from a variety of data source: federal and state agency, PSC (coded-wire tag), and tribal. The NOAA Salmon Population Summary describes the method.

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.

CSS Response: The ^ symbol for predicted distinguishes between predicted and observed. Lower case s should have been for survival. Upper case S is still spawners. Table 2.1 distinguishes between state variables and parameters, and between estimated parameters and derived parameters. If a parameter gets its value from an equation, it is a derived parameter, otherwise it's estimated.

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?

CSS Response: Uninformative priors were used in the MCMC simulations.

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)?

CSS Response: Walters et al. (2013) estimated the global smolt per redd Beverton Holt productivity parameter to be 275, which would correspond to 137.5, 92, and 69 smolts per spawner at 2, 3, and 4 spawners per redd respectively (the range of spawners per redd observed in the Grande Ronde population). This study included predicted juvenile production dating back to farther than the Walters et al. study, which might explain the higher productivities estimated in this analysis.

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

CSS Response: Three survivals were entirely derived from environmental data (see Equations 2.10–2.12). They are shown in Figures 2.5 and 2.6 to illustrate the relationships between survival and 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?

CSS Response: The maximum likelihood approach we used uses a non-linear optimization algorithm that finds the mode of the likelihood surface and evaluates the variability in the likelihood near the mode with respect to each estimated parameter. The partial derivatives of the likelihood with respect to each parameter are estimates of the variance in parameter estimates. Given that this information is available within the algorithm we used, the standard deviation is reported. It is not Bayesian, but it is a valid representation of the variance in the estimate.

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.

CSS Response: The deviates we estimated were not random variables. Had a state-space model been implemented, the deviates could have been treated as random effects and a likelihood component could have been added to explain that variation, in which case only variance parameters would have been estimated. But in this case, it was implemented as maximum likelihood estimation with 43 deviates estimated.

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.

CSS Response: Figures 2.5 and 2.6 are correlation plots. They provide a visual representation of relationships predicted by Equations 2.10–2.12 (columns 7, 8, and 9) in relation to environmental variables and survivals estimated from PIT tags.

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

CSS Response: It's not clear how Figure 2.10 would be BLUPs. They are the dependent variables in a prediction made with parameters estimated from maximizing likelihood.

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?

CSS Response: Maximum likelihood estimates of parameters occur at the mode of the likelihood (i.e., the highest point on the likelihood surface). The mode of the sample of the posterior of a MCMC simulation is not necessarily the MLE at the mode of the likelihood. We include the mean and standard deviation of the posterior sample for reference against the MLE results.

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.

CSS Response: Some estimated parameters were bounded within reasonable ranges, but in some cases there was insufficient information to provide an estimate. In the case of (b), the capacity parameter, only low abundances are seen in the data, so the capacity might as well be infinite, and the most likely value is near the upper bound. In the case of (q), the interpretation would be that the early ocean survival rate of transported fish in the absence of the PDO and UPW effects is at the lower range of bounds placed on that parameter. Since that corresponds to approximately 5% survival, the result implies that the rate fluctuates around 5% depending on the variation in PDO and UPW. The PDO and UPW effect is shared with fish that migrated through the hydro system, and it might be possible that PDO and UPW do not affect transported and in-river fish the same way, but we made the assumption that both groups were affected by those variables in the same way. We made that assumption to keep the number of parameters to a minimum.

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.

CSS Response: There are several reasons why we have chosen to conduct our analyses of associations between fish travel time, instantaneous mortality rate, and survival outside of the CJS mark-recapture model (e.g., the MARK program).

First, we have followed previous ISAB suggestions that "an interpretation of the patterns observed in the relation between reach survival and travel time or flow requires an understanding of the relation between reach survival, instantaneous mortality, migration speed, and flow" (ISAB 2003). Towards this end, we have conducted extensive analyses on the factors that influence migration speed (i.e., fish travel time). The models that we have developed, including mixed-effects models, have been highly successful at improving understanding of the factors that influence migration speed. Also following previous ISAB suggestions, we have developed estimates of instantaneous mortality rates and models for improving understanding of the factors that influence instantaneous mortality rates. We are not aware of any way to estimate instantaneous mortality within the MARK program or any way to analyze covariates relative to instantaneous mortality within the MARK program. The models for fish travel time and instantaneous mortality in turn are used to improve understanding of reach survival through the framework provided by the exponential survival model. We have led the efforts to improve understanding of the relationships between "reach survival, instantaneous mortality, migration speed, and flow" (ISAB 2003), yet despite the success of these efforts the ISAB has chosen to pejoratively characterize these efforts as a "statistics-on-statistics" approach. This characterization is unjustified and inappropriate, especially when it was the ISAB who originally made suggestions to work on the relationships between reach survival, instantaneous mortality, and migration speed.

It is possible to incorporate covariates directly into the CJS mark-recapture model when estimating and analyzing patterns of survival. However, there are several complicating issues to consider with such an approach. First, each sub-reach requires estimation of a parameter for the effect of each covariate under consideration. For example, the LGR-MCN reach consists of four sub-reaches (LGR-LGS, LGS-LMN, LMN-IHR, and IHR-MCN) and would require estimation of a coefficient for the effect of each covariate in each reach. There are also potential delayed effects of covariates in previous reaches influencing survival in subsequent reaches. Because of this, a fully-parameterized, four sub-reach model with delayed effects would require estimation of up to nine coefficients for each covariate, increasing model complexity. In contrast, analyzing covariate effects on the product of the four sub-reaches could be accomplished by using just one estimated coefficient when analyzing survival outside the CJS model. Another issue is the frequent observation of oscillations among adjacent reaches where survival estimates are unusually high in one reach followed by an unusually low estimate in the next reach (or vice versa), but the product of the reaches is typically reasonable. The multiplication of several sub-reach survival estimates in order to estimate the full reach survival rate helps to stabilize the variability that occurs within the sub-reach survival estimates. Finally, it does not appear that the MARK program allows for the investigation of covariate effects within complex correlation structures, such as mixed-effects models. MARK can account for relatively simple random effects, but the level of complexity that can be analyzed (e.g., random intercepts and random slopes on covariates) is very limited.

Analyzing the effects of environmental or management covariates on survival outside of the CJS model is a common, accepted practice within the peer-reviewed scientific literature. An abbreviated list of peer-reviewed, mark-recapture studies on Pacific salmon where covariate effects on survival are analyzed outside the CJS model include: Briscoe et al. 2005, Clemens et al. 2009, Connor et al. 2003, Connor and Tiffan 2012, Griswold et al. 2011, Hockersmith et al. 2003, Hoffnagle et al. 2006, Paulsen and Fisher 2001, Pess et al. 2011, Roni et al. 2012, Smith et al. 2002, and Smith et al. 2003. Clearly, analyzing covariate effects on survival outside the CJS model is an accepted analytical approach.

Some key questions to consider when evaluating alternative analytical approaches are whether the approach properly accounts for uncertainty, whether the approach introduces biases in some way, and whether the approach is capable of providing an adequate characterization of complex biological phenomena. We believe that the analyses presented in Chapter 3 are successful on each of these fronts. The analyses properly account for uncertainty in the underlying data through the use of the estimation of uncertainty in response variables, the use of weighted regressions, and the applications of mixed-effect model structures that account for correlations among observations. The analyses do not appear to result in systematically biased predictions relative to observations. The analyses are successful in characterizing the variability in fish travel time, instantaneous mortality, and survival rates using parsimonious models firmly grounded in empirical observations. These factors deserve careful consideration prior to disparaging the analyses in Chapter 3 as an inferior, inadequate “statistics-on-statistics” approach.

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.

CSS Response: Table 3.1 presents the table summarizing the locations, species, rearing type (i.e., hatchery or wild), and life-history type (i.e., yearling or subyearling Chinook salmon) that were analyzed.

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.

CSS Response: Additional text was added to clarify the rationale behind the 7-day windows. Briefly, the rationale behind using the 7-day window around the median passage date is to develop an index of exposure to the environmental variables analyzed (e.g., spill, water transit time, temperature) that aligns with the timing of smolt passage at each dam. This choice is not arbitrary, as the vast majority of smolts pass during these 7-day windows around the median passage date and experience the spill, temperature, and water transit times that occur within these windows. Similar indices have been developed and applied by other researchers (e.g., Smith et al. 2002, Connor et al. 2003). The regression models developed using these indices explain a high degree of the variation in fish travel time, instantaneous mortality rates, and survival rates (Table 3.3), which supports the adequacy of the indices. We are currently developing additional indices for spill that estimate total spillway passage at each dam, and plan to incorporate these indices in future reports.

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.

CSS Response: We believe the coefficients of determination (not variation) values presented in Table 3.3 are appropriate measures of model fit, reflecting the proportion of the overall variability that is accounted-for by the models. The model-averaged predictions that were used to calculate the coefficients of determination are, by definition, generated by the lowest-AIC_C set of models, which is another measure of model fit. We have included some discussion on why the models for FTT generally fit better than those for Z and survival, along with the lower coefficients of determination values for yearling Chinook.

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.

CSS Response: We added text explaining that we conducted simulations of the predicted fish travel times and survival rates that would have occurred if spill levels had remained high at Ice Harbor Dam. Those simulations indicate that fish travel times would have been 0.94 d faster and survival rates would have been 2.3% higher if spill levels had remained at 75% over the 2005–2013 period.

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?

CSS Response: Analyses conducted by Skalski et al. (1998) found that detection probabilities were not affected by upstream capture histories and that the test of homogeneity (TEST 3) proposed by Burnham et al. (1987) did not indicate significantly different capture probabilities for detected and non-detected fish at Lower Granite Dam. These and other analyses support the assumption that detected fish (i.e., fish entering the bypass system) are a random sample of the population of individuals passing the dam.

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?

CSS Response: The table value of 39% is correct and the text value was corrected to also be 39%.

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

CSS Response: We have added references supporting the assumption that S tends to be 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.

CSS Response: All models include an intercept term, therefore the coefficient of determination is a valid measure of the proportion of variance explained.

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?

CSS Response: The R^2 values are calculated as the squared Pearson correlation coefficient between the back-transformed predictions and the observed values.

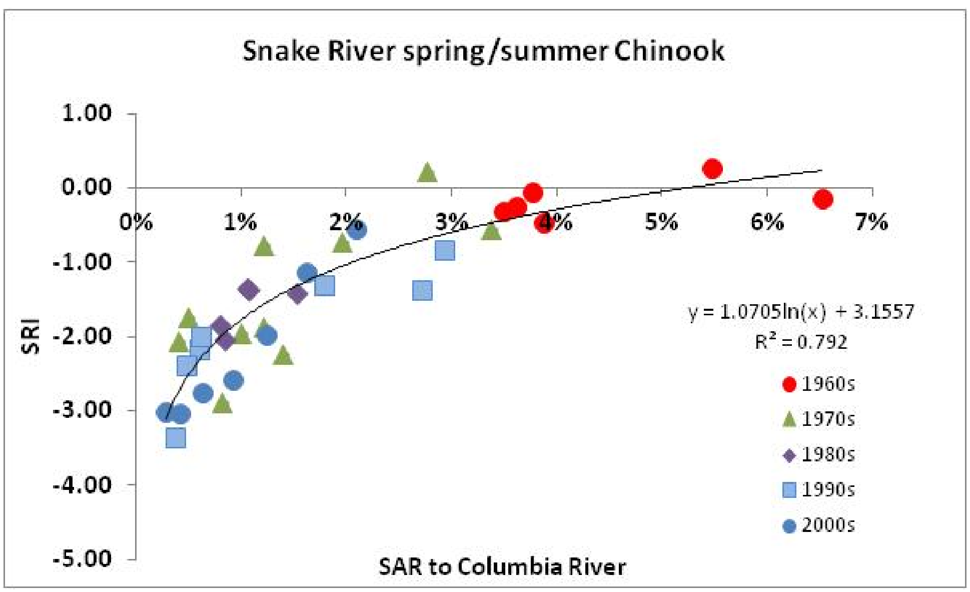
Chapter 4. Patterns in Annual Overall SARs

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.

CSS Response: The emphasis of this section to date has been to examine the level of SARs needed to achieve population replacement given recent abundance and freshwater survival. Life cycle survival rate is the product of the life stage survival rates, any one of which may influence (and correlate with) adult recruits/spawner. The primary inference from the strong correlation observed between adult recruits/spawner and SAR would be that the smolt-to-adult life stage is very important to overall life cycle survival. These observations support earlier conclusions of Petrosky et al. (2001), who found that "the decline of Chinook populations following completion of the hydrosystem was not accompanied by major survival changes in the [freshwater spawning and rearing] FSR life stage. FSR productivity showed no significant decline, and the FSR survival rate decline was small relative to the overall decline. However, significant survival declines did occur in the smolt-to-adult stage coincident primarily with hydrosystem completion, combined with poorer climate conditions and possibly hatchery effects." Study results also indicate that potential survival rate improvements from tributary habitat restoration fall short of the minimum survival rate improvements to meet viability for the majority of MPGs across the Snake River spring/summer Chinook ESU (Budy and Schaller 2007). Note also that adult recruits in the numerator of both quantities (adult returns/spawner and SAR) in this section are from different data sets. Adult recruits in the productivity estimates derive from run reconstruction of spawning ground surveys, while the adult recruits in the SAR estimates represent PIT-tag detections at Lower Granite Dam.

The authors recognize the need to incorporate density dependence into the examination of population productivity versus SAR (p. 100).

CSS Response: We have begun preliminary work on this topic, using the recruitment functions from Schaller et al. (2014). Schaller et al. defined survival rate indices (SRIs) as the deviation of the observed ln(R/S) from the pre-1970 expected ln(R/S), where R represents pre-harvest recruits to the Columbia River mouth. A plot of the average SRI from 18 Snake River populations versus SARs to the Columbia River mouth is shown below. SARs explained a large portion of the variability in recruitment during this time period (1964–2007 smolt migration years), after accounting for density dependence in life cycle survival rates derived from ln(R/S).



Comparison of survival rate indices (SRIs) and SARs of Snake River spring/summer Chinook, 1964-2007 smolt migration years. SRIs represent the annual average from 18 Snake River populations, where SRI = 0 is the average expected productivity for the pre-1970 brood years (Schaller et al. 2014). SARs to the Columbia River mouth are from Table 4.63 and the 2012 CSS annual report (Tuomikoski et al. 2012). The prediction line is fitted through all years of data.

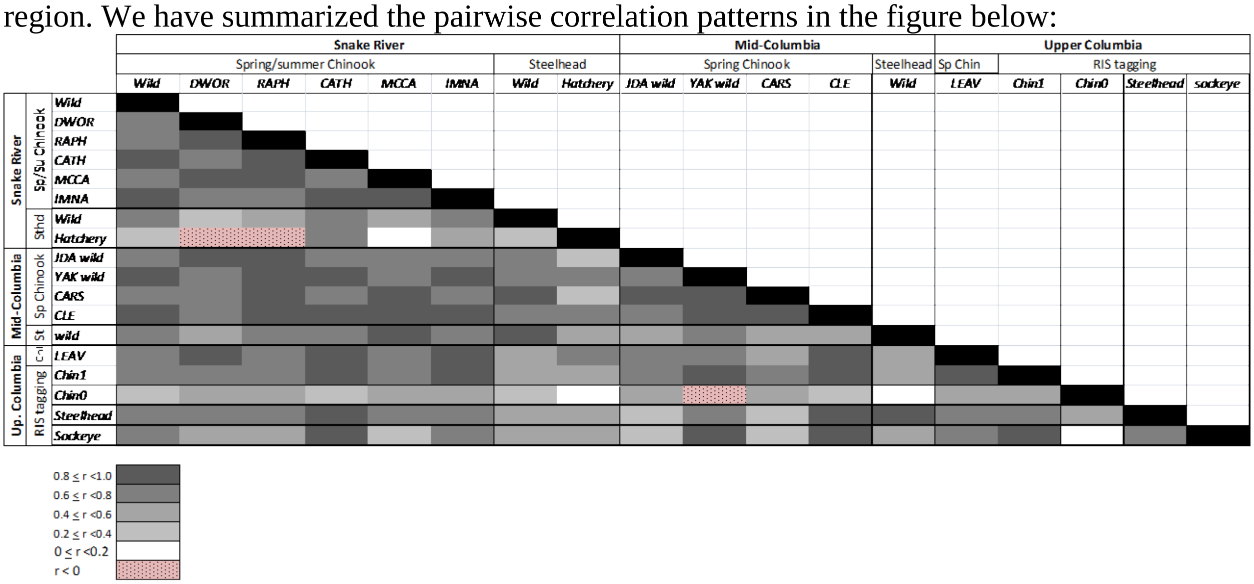
They should also examine the level of smolts per spawner needed to achieve viability given current SAR values while also considering the observed density dependent relationship between smolts per spawner and spawners.

CSS Response: This will be possible in future analyses. Such analyses should also consider the feasibility of survival rate increases in both freshwater spawning and rearing (FSR) and SAR life stages (e.g., Budy and Schaller 2007; UCSRB 2014; Haeseker et al. 2012; Hall and Marmorek 2013). We would note that there is little room for improvement in FSR productivity for some populations, such as those in the Middle Fork Salmon MPG. The Chapter 2 modeling is being developed specifically to examine changes in both SAR and FSR life stages for the Grande Ronde/Imnaha MPG, which contains some population segments with substantial potential to improve FSR habitat.

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).

CSS Response: We can investigate this in the future, although we note differences in geographic scale of data (and definition of regions) between the CSS and Pyper et al. (2001). Pyper et al. (2001) found strong evidence of positive co-variation among pink salmon stocks within region and between certain adjacent regions, but no evidence of co-variation between stocks of distant regions (separated by 1000 km) of the Northeast Pacific. (Note: all Columbia River groups used in the CSS would be considered to be within a single region using the geographic scale of Pyper et al.).

The CSS now has ten or more years of SAR estimates for 18 groups of salmon and steelhead in the interior Columbia. Groups with a minimum of ten years data include six spring/summer Chinook groups and two steelhead groups in the Snake River region, four spring Chinook groups and one steelhead group in the Mid-Columbia region, and two spring Chinook groups, one summer/fall Chinook group, one steelhead group and one sockeye group in the upper Columbia region. We have summarized the pairwise correlation patterns in the figure below:



Correlation of SARs of Snake River, Mid-Columbia and Upper Columbia salmon and steelhead groups, 1997–2012 smolt migration years. Groups included had a minimum of 10 years data. SARs represent smolts at uppermost dam and adult plus jack (for salmon) returns to uppermost dam with PIT-tag detection capabilities. Snake River SAR data are from Tables 4.1, 4.15, 4.17, 4.19, 4.21, 4.23, 4.33 and 4.40, respectively. Mid-Columbia SAR data are from Tables 4.42, 4.43, 4.45, 4.47 and 4.51, respectively. Upper Columbia SAR data are from Tables 4.54, 4.59, 4.60, 4.61 and 4.62, respectively.

The general pattern in the CSS time series is for strong, positive co-variation of SARs within species and across regions of the Interior Columbia for stream-type Chinook and steelhead. There were only three negative pairwise correlations between groups (range -0.02 to -0.09). Pairwise correlations within and between spring/summer Chinook and spring Chinook regional groups were consistently strong (average r = 0.80, range 0.61 to 0.94). Pairwise correlations within and between steelhead regional groups were moderately strong (average r = 0.55, range 0.33 to 0.80). Subyearling Chinook tagged by the SMP at Rock Island Dam showed weaker correlation with other species/run types (average 0.34, range -0.04 to 0.56). As we noted in the Chapter 4 discussion, a high degree of inter-regional correlation indicates common environmental factors are influencing survival rates from outmigration to the estuary and ocean environments.

This correlation analysis is consistent with results of multivariate analysis of Columbia River salmon populations that indicate similar ocean and river variables explain variation in survival rate patterns (Schaller and Petrosky 2007; Scheuerell et al. 2009; Petrosky and Schaller 2010; Haeseker et al. 2012; Schaller et al. 2014). The additional SAR time series being developed in CSS could contribute to a broader analysis of these similarities and differences in patterns across Columbia River population groupings.

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.

CSS Response: A description of the bootstrap procedure is available in Chapter 4 of the CSS 2002 Annual Report (Berggren et al. 2002) available online at http://www.fpc.org/documents/CSS/final_2002_CSS_AnnualReport.pdf. The bootstrap procedure resamples PIT-tagged fish with replacement using the methods described by Efron and Tibshirani (1993). The unit of resampling is the individual fish and all its detection information. In particular, information about fish detection history is resampled, including each dam where fish were detected as juveniles, as well as the location fish may have been removed for transport. Additional data for each fish would include whether a fish was detected at the estuary trawl, or any adult detections at various dams. The bootstrap procedures are used to resample PIT-tagged fish in order to provide non-parametric confidence intervals for calculated parameters of interest. The bootstrap methods themselves did not change after 2006 (i.e., individual PIT-tagged fish are still resampled with replacement the same way as prior to 2006). The method for calculating the overall SARs in particular did change in 2006. The overall SAR computation became much more straight forward in 2006, when we began to pre-assign tagged fish to be treated the same as unmarked fish migrating through the Snake and Columbia River hydro projects. We inserted a reference to the CSS 2002 annual report (Berggren et al. 2002) bootstrapping description into the Chapter 4 methods section.

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?

CSS Response: See answer above. Because the unit of resampling is the individual fish, both juvenile detection history and adult return information are resampled simultaneously. Adult data represent counts of adult detections at dams, not estimates. Regardless of the method of determining adult numbers, because of the resampling of all information for individual PIT-tagged fish, the adult data associated with each fish are resampled in each bootstrap run. The variability inherent in the resampling procedure captures the uncertainty in adult return numbers as well as juvenile fish detection.

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.

CSS Response: The most appropriate units for both R/S and SAR depend on the questions being asked. Recruitment to the Columbia River mouth is used to address pre-harvest productivity, and is particularly relevant over time periods when harvest rates have varied (as in the Chapter 2 time series). Recruitment to spawning grounds is most relevant to questions of viability (e.g., ICTRT 2007) and spawning escapement objectives. We attempted to clarify this in the final Chapter 4, and added the notation Rsg for spawning grounds recruits.

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

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

CSS Response: We added the following discussion of low survival results displayed in Table 5.9. "***Survival estimates for wild subyearling fall Chinook are shown in Table 5.9. Estimates for wild subyearling fall Chinook were generally much lower than those estimated for hatchery released fish. However, caution should be used when evaluating the estimates of reach survival for wild subyearling fall Chinook, since they are based on small sample sizes and relatively large expansions (either 51% or 77%) were used to derive estimates of full reach survival. Given that difficulty in interpreting these estimates, there could be several possible explanations as to why wild fish exhibit lower reach survivals that hatchery origin fish. Size at emigration and rate of emigration are likely smaller and later for wild fish. Wild marked fish tend to be released at sizes 10mm to 20mm smaller than hatchery origin fish. In addition, Connor et al 2013 suggest that density-dependence in recent years may affect the behaviour of these fish after marking, making them more likely to emigrate at a smaller size. There is also the possibility that hatchery supplementation has altered the genetic composition of the wild population and this in turn has led to the change in behaviour observed by Connor et al. Whatever the cause, if fish are emigrating at a smaller size, that could explain some of the difference in survival between hatchery and wild mark groups observed in CSS estimates."***

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*

CSS Response: We added the following sentence to the text to highlight relationship between SARs with and without jacks. "***SARs that included jacks were consistently higher than those excluding jacks except for migration year 2007, when SARs for all groups jacks or not were very low (see Tables 5.10 to 5.13).***" In regard to the relationship between wild subyearling Chinook survival and SARs, it must be noted that reach survival estimates for wild subyearling fall Chinook were based on relatively large expansions in all years for which SARs were reported (either 51% or 77% expansion). We would use caution in interpreting the relationships between those reach survivals and SARs relative to reach survivals and SARs of hatchery groups that often had much higher LGR populations from which reach survivals were based, and often had only single reach expansions. In general, we would agree however, that it does appear that SARs for wild fish were relatively high compared to hatchery groups. We would also agree that conclusions about wild subyearling Chinook TIRs would be very tentative at this time.

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.

CSS Response: We added summary information regarding *D* values in paragraphs that previously only described patterns in TIRs. *D* estimates for fall Chinook tended to be nearly all uniformly below 1 for the years included in the CSS report. This pattern seemed relatively uninformative. That differed from yearling Chinook and steelhead, which prior to 2006, had shown *D* values above 1 for several years. We added a concluding paragraph about *D* that was stated as follows: **"Estimated *D* values for subyearling Snake River fall Chinook were below 1, for nearly all groups in the years 2006 to 2011. That was similar to patterns seen in yearling Chinook and steelhead (hatchery and wild groups) in the same years. A longer time series for subyearling Chinook would be helpful to determine if *D* estimates would have been higher prior to 2005 (beginning of court ordered summer spill) similar to the pattern seen for hatchery and wild steelhead groups that had *D* values that were well above 1 for several years."**

"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?"

CSS Response: We have clarified the language in the report so that the observation is understood within the context of the CSS monitoring program and the operation of the FCRPS. The smolt transportation program was developed and implemented to mitigate for the recognized adverse impacts of the development and operation of the FCRPS on salmon and steelhead migrating through the mainstem Snake and Columbia rivers. The recognition that the development and operation of the FCRPS was resulting in significant salmon and steelhead mortalities resulted in the implementation of the smolt transportation program to mitigate these adverse impacts. The language has been revised to ***"Based on TIRs of adult returns to LGR it***

appears that the juvenile smolt transportation program does not mitigate for the adverse impacts of the operation of the FCRPS on fall Chinook."

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?

CSS Response: The models actually were effective at predicting holdovers but not in the way we anticipated. We anticipated being able to predict holdovers based on size and release date in such a way that we would be able to identify and remove certain individual fish with high holdover probability from within a particular release group. What we found was that all individuals within a release group usually had similar size and release date so that effectively we identified an entire release group as having high holdover probability. This made sense in terms of the data but did not allow us to remove individual fish from release groups since it really identified entire release groups that were likely to holdover. The model was effective, but not for our intended purpose.

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.

CSS Response: We modified the comments to better reflect the apparently higher SARs in 2008 and the small number of groups with SAR data in 2007 as follows: "**Patterns in overall SARs were similar for several groups, with the highest return year being 2008 (Figure 5.2). For those groups that had PIT-tag marking in 2007, the year when the marking study was cancelled, SARs were lowest. Lower SARs were also seen 2009 for nearly all groups as well. The relatively high return rate seen in 2008 is similar to what was seen for many Snake River spring/summer Chinook mark groups, which had their highest returns of recent years, in 2008. We anticipate that when 3-salt returns from 2011 have been included, that SARs from that year could exceed 2008 for most groups. Those data will be available in the 2015 report.**"

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

CSS Response: See comment above for detailed response. Greater details were added to include discussions and conclusions about *D*.

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.

CSS Response: We have made efforts to collect data in order to answer questions on topics (a, b, and c) listed above. On topic b, adults in excess of brood stock needs are removed from the hatchery prior to spawning, and for these fish sex determination is technically and logistically difficult. However, sex is determined for the vast majority of the adult returns.

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.

CSS Response: The current study represents a realistic scenario for hatchery CWT and PIT tagging studies. Hatchery spring Chinook salmon are typically tagged weeks to months prior to release and the study design allows for the estimation of tag loss/mortality from the time of tagging until release for PIT-tagged fish. Preliminary results indicate that immediate tag loss/mortality rates are low during the time in the hatchery prior to release. The size of the fish is typical for hatchery spring Chinook populations. Given the sample sizes required, it would be extremely difficult to conduct a parallel study on 60 mm migrating smolts captured in the field during warm-water periods. However, a study on hatchery fall Chinook could be performed to evaluate the effects of small smolt size on tag loss/mortality.

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.*

CSS Response: We have added some discussion on potential reasons for the lower return rates of dual-tagged fish.

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.

CSS Response: The original sample size calculations (Brignon and Haeseker 2011) were conducted using an ANOVA model with $\log_e$(SAR) as the dependent variable and with inverse-variance weighting because empirical (Peterman 1981) and theoretical (Hilborn and Walters 1992) analyses support the assumption that survival rates (e.g., SARs) tend to be log-normally distributed. A logistic or Poisson model would be more appropriate if we were to analyze the number of adult returns rather than smolt-to-adult return rates. We will investigate which type of model best meets model assumptions when it comes time to conduct those analyses and may conduct simulations to see whether model type influences statistical power.

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?

CSS Response: The statement of no evidence of a difference was based on visual inspection of the survival estimate plot, which showed overlapping confidence intervals. We have changed the text to clarify this point.

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.

CSS Response: The text was clarified to better describe the rearing procedures that were used.

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?

CSS Response: The rationale behind the sample sizes are described in Brignon and Haeseker (2011). A release of 75,000 individuals in the CWT-only group combined with 15,000 individuals in each of the PIT-only and dual-tagged groups achieved a power of nearly 80% to detect a PIT-tag effect on SARs. The 75,000 CWT-only individuals were economically feasible and were consistent with hatchery objectives for their CWT program. Additional PIT tags would have improved power, but funding was only available to tag 15,000 individuals in each of the PIT-only and dual-tagged groups.

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

CSS Response: The text was clarified to better describe the rearing procedures that were used.

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?

CSS Response: The text was clarified to better describe the release procedures that were used.

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?

CSS Response: The text was clarified to better describe the reasons why it was necessary to euthanize dual-tagged fish.

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.

CSS Response: The power analyses were focused on the primary question of whether there are differences in the SARs of each of the three tag groups. Tagged fish mortality or shedding during rearing and poor juvenile detection efficiency at release will not impact these analyses because we are using the number of smolts that were originally tagged as the starting population for calculating SARs. However, the retention samples provide data to also correct the SARs for tag retention, and those estimates can be calculated. We have added text to clarify how smolts were enumerated during rearing. In the future we do plan on making comparisons between the retention sample estimates of shedding/mortality and the estimates of shedding/mortality for PIT-tagged fish derived using the release channel array.

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.

CSS Response: The text was clarified to better describe the release procedures that were used.

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.

CSS Response: Juvenile survival rates were calculated using the same method (i.e., the CJS model), but the survival rates were separately estimated for the PIT-only and dual-tagged group to assess whether there were differences in juvenile survival between the two groups. The text was changed to better clarify this point.

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.

CSS Response: We have provided a table with the number of adult returns by age and by tag group. We have also provided a table of tag loss for the dual-tagged group, the only group that can be used to assess tag loss in adults. In subsequent reports we plan on assessing tag loss in adults after hatchery entry.

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.

CSS Response: We have added some discussion of potential reasons for low preliminary returns of the dual-tagged group.

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.

CSS Response: No adjustment, such as a Bonferroni correction, was made for multiple testing. It is true that without correcting the 90% CI, about 10% of these tests will find evidence of differential delayed mortality even if it does not exist (type-I error). However, adjusting the type-I error rate for multiple comparisons will also inflate the type-II error (Rothman 1990). That is, the number of times that we fail to detect differential delayed mortality effects, when in fact it exists, also increases. For certain groups where *D* estimates are available from 1994 – 2013, accounting for 20 comparisons would result in adjusting the overall type-I error to 0.1/20 = 0.005 (i.e. Bonferonni correction) or equivalently computing a 99.5% CI for *D*. This would yield unwieldy wide CIs that would not be sensitive enough to detect differential delayed mortality effects. A more pragmatic approach is simply to recognize that given 20 such tests at a type-1 error rate = 0.1, we would expect to observe 2 significant results due to chance alone. For instance, when examining Table A.23 which summarizes TIR and *D* for wild Chinook, 13 significant results for D were observed. These results still provide strong evidence for differential delayed mortality as we would expect only 2 to arise due to chance.

More recently developed statistical techniques, such as Bayesian hierarchical models, are arguably immune to the problem of multiple comparisons (Gelman et al. 2012). Assuming that the observed TIRs or *D*s all came from the same hierarchical distribution would in effect slightly shift their values closer together. This would make comparisons more conservative. In this paradigm, all of the parameters would be represented in one coherent model and the perceived problem of multiple comparisons would be built in from the start. Given the large number of TIR and *D* statistics that are reported for different species and stocks, a series of hierarchical Bayesian models is probably not practical in this setting, but could be pursued for a more focused application.

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

Response: For a given migration year, all hatchery steelhead are combined into a single group. A single bootstrap run is then done for this large group.

V. Editorial Comments

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

CSS Response: NPH was taken from Petrosky and Schaller (2010) and was calculated as:

$$\mathrm{N\_Power\,House}(i) = \mathrm{Ndams}(i) - \sum_{j=1}^{\mathrm{Ndams}(i)} \mathrm{Prop\_spill}(i,j),$$

where Ndams is number of dams completed in year *i* and Prop_spill is the spill proportion at dam *j* in year *i*. The method assumed spill efficiency of 1:1.

PITPH was developed as an estimate of the probability of fish encountering a powerhouse, or multiple powerhouses, given a model of fish passage (i.e., detection probability and spill passage efficiency) based on PIT-tag detection data. Both measure powerhouse passage, but PITPH is derived from PIT-tag data, and used dam-specific models of passage efficiency based on flow, spill and presence of surface passage structures to determine the proportion of fish that would pass through the powerhouse.

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.

CSS Response: We are working to remove redundant abbreviations from the document.

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."

CSS Response: Survival rate will be replaced with survival probability where appropriate.

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

CSS Response: We made this change.

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.

CSS Response: We updated several figures in the document with black and white contrasting symbols and lines in mind.

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.

CSS Response: Replaced the sentence with the following: **"This method is used to estimate a population of PIT-tagged smolts alive in the tailrace of Lower Granite Dam and estimate their survival through the hydrosystem."**

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?

CSS Response: We rewrote and reorganized much of the section referred to in order to consolidate ideas and information.

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.*

CSS Response: We rewrote and reorganized much of the section referred to in order to consolidate ideas and information.

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.

CSS Response: Thanks. We will work towards adding more clarity to 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.

CSS Response: Essentially, TIR is a ratio of LGR to LGR transport SAR divided by LGR to LGR in-river SAR. While *D* is the ratio of BON to LGR transport SAR divided by the BON to LGR in-river SAR. Unfortunately, both ratios involve data from the entire life cycle of fish so we are struggling to decide what type of a diagram would provide greater clarity than the equations that describe them.

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?

CSS Response: Data were arranged by funding/tagging agencies and species.

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

CSS Response: We appreciate this comment, and normally we would present percentages on a 0 to 100% scale, but in doing so the figure became too small to view clearly. So we decided to maintain the figure as is.

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.

CSS Response: The subtlety of this distinction is clarified in the revised chapter.

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

CSS Response: PITPH and WTT were added to determine if variation in mainstem survival could be explained by water velocity and powerhouse contact rate.

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.

CSS Response: Clarification provided in the revised chapter.

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

CSS Response: Corrected.

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

CSS Response: LCH builds on LCX model, further separating juvenile migration survival into in-river and transported components.

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?

CSS Response: Further clarification provided throughout the revised chapter.

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

CSS Response: Further clarification of process and observation error model distinctions provided throughout the revised chapter.

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.

CSS Response: Corrected to year-salt notation.

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?*

CSS Response: Yes, we computed c-hat using the ratio of the deviance to its degrees of freedom. Text was added to clarify this.

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.

CSS Response: Hats added to the random variables.

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

CSS Response: Hats added to the random variables.

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).

CSS Response: As was mentioned on page 60, Box-Cox power transformations indicated that the square-root-transformation was most supported across the ten species-reach combinations that have been analyzed.

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?

CSS Response: The text reports juvenile survival rate as a proportion, consistent with Figure 4.14. The confidence limits exceed 1 in a few cases for the 2-week cohorts due to small sample size and limited juvenile detection capability.

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

CSS Response: We revised units in the text to proportions for consistency.

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?

CSS Response: We revised the notation from ln(S/S) to ln(Rsg/S) to clarify that these are recruits to the spawning grounds as defined on page 78 of the draft report.

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?).

CSS Response: In our opinion the number would have 3 to 4 zeros in front of them making them hard to read as well as difficult to understand in terms of the SARs reported. We retained the presentation format from the draft.

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

CSS Response: We stand by the usage by which we meant "…to place side by side; or juxtapose" for comparison.

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

CSS Response: We believe that the usage was correct "…the quality of being concise: brevity." We used concision to be concise, although the origin of the word gave us pause.

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?

CSS Response: The power analyses (Brignon and Haeseker 2011) assumed that CWT-only fish had an SAR that was 25% higher than dual-tagged fish (i.e., the assumed magnitude of the PIT effect), which is a slightly smaller detectable effect size than the PIT-tag effect estimated by Knudsen et al. (2009).

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

CSS Response: Text changed.

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

CSS Response: Table captions edited to define NA as Not Applicable.

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.

CSS Response: The table caption and column heading have been changed to reflect the total numbers that were tagged by group.

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

CSS Response: A table has been substituted to present the data plotted in Figure 6.2.

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

CSS Response: We have provided a table as a substitute for Figure 6.4. Being color-blind myself, I understand the challenges of deciphering color.

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.

CSS Response: We considered the ISAB's comment for this figure and corrected many figures throughout this Appendix to avoid using only color to distinguish between separate lines. Separate lines are now distinguishable by color and/or symbols.

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

CSS Response: Correction made.

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

CSS Response: We have redesigned this figure to make it easier to distinguish different species to make it more conducive to photocopying.

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

CSS Response: We have redesigned this figure to make it easier to distinguish species, display 2001, and display those years that represent the Court Ordered Spill program (2006–2011).

VI. References (provided by ISAB)

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.

CSS Response References

Berggren, T., H. Franzoni, L. Basham, P. Wilson, H. Schaller, C. Petrosky, E. Weber, R. Boyce and N. Bouwes. 2002. Comparative Survival Study (CSS) of PIT tagged spring/summer chinook. 2002 Annual Report, Migration Years 1997 – 2000, Mark/Recapture Activities and Bootstrap Analysis. BPA Contract #19960200. Available at http://www.fpc.org/

Brignon, W.R and S.L. Haeseker. 2011. PIT-tag effects on hatchery salmonids: Carson National Fish Hatchery spring Chinook salmon. Annual Report 2011 and Work Plan 2012. Columbia River Fishery Program Office. U.S. Fish and Wildlife Service. 8 pp. (Available at http://www.fws.gov/columbiariver/publications.html)

Briscoe, R.J., M.D. Adkison, A. Wertheimer, and S.G. Taylor. 2005. Biophysical factors associated with the marine survival of Auke Creek, Alaska, coho salmon. Transactions of the American Fisheries Society 134: 817-828.

Budy, P. and H. Schaller. 2007. Evaluating tributary restoration potential for Pacific salmon recovery. Ecological Applications 17: 1068-1086.

Burnham, K. P., D. R. Anderson, G. C. White, C. Brownie, and K. H. Pollock. 1987. Design and analysis methods for fish survival experiments based on release-recapture. American Fisheries Society Monograph 5. Bethesda, MD. 437 pp.

Clemens, B.J., S.P. Clements, M.D. Karnowski, D.B. Jepsen, A.I. Gitelman, C.B. Schreck. 2009. Effects of transportation and other factors on survival estimates of juvenile salmonids in the unimpounded lower Columbia River. Transactions of the American Fisheries Society 138: 169-188.

Connor, W.P., H.L. Burge, J.R. Yearsley, and T.C. Bjornn. 2003. Influence of flow and temperature on survival of wild subyearling fall Chinook salmon in the Snake River. North American Journal of Fisheries Management 23: 362-375.

Connor, W.P., and K.F. Tiffan. 2012. Evidence for parr growth as a factor affecting parr-to-smolt survival. 2012. Transactions of the American Fisheries Society 141: 1207-1218.

Connor, W. P., Tiffan, K. F., Plumb, J. M., & Moffitt, C. M. (2013). Evidence for density-dependent changes in growth, downstream movement, and size of Chinook salmon subyearlings in a large-river landscape. *Transactions of the American Fisheries Society*, *142*(5), 1453-1468.

Efron, B. and R.J. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, NY.

Gelman A, Hill J, Yajima M. 2012. Why we (usually) don’t have to worry about multiple comparisons. Journal of Research on Educational Effectiveness 5:189- 211

Griswold, R.G., A.E. Kohler, and D. Taki. 2011. Survival of endangered Snake River sockeye salmon smolts from three Idaho lakes: relationships with parr size at release, parr growth rate, smolt Size, discharge, and travel time. North American Journal of Fisheries Management 31: 813-825.

Haeseker, S.L., J.A. McCann, J. Tuomikoski and B. Chockley. 2012 Assessing freshwater and marine environmental influences on life-stage-specific survival rates of Snake River spring-summer Chinook salmon and steelhead. Transactions of the American Fisheries Society 141:121-138.

Hall, A. and D. Marmorek. 2013. Comparative Survival Study (CSS) 2013 Workshop report. Prepared by ESSA Technologies Ltd., Vancouver, B.C. for the Fish Passage Center (Portland OR) and U.S. Fish and Wildlife Service (Vancouver, WA). 47 p. plus Appendices.

Hilborn, R. and C. J. Walters. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Chapman and Hall, New York.

Hockersmith, E.E., W.D. Muir, S.G. Smith, B.P. Sandford, R.W. Perry, N.S. Adams, and D.W. Rondorf. 2003. Comparison of migration rate and survival between radio-tagged and PIT-tagged migrant yearling Chinook salmon in the Snake and Columbia Rivers. North American Journal of Fisheries Management 23: 404-413.

Hoffnagle, T.L., R.W. Carmichael, and P.J. Keniry. 2006. The effect of moderately increased and variable raceway flow rates on juvenile physiology, survival, and adult return of hatchery-reared Chinook salmon. Transactions of the American Fisheries Society 135: 1567-1577.

ICTRT (Interior Columbia Technical Recovery Team). 2007. Viability criteria for application to Interior Columbia Basin Salmonid ESUs. Technical Review Draft March 2007. Interior Columbia Basin Technical Recovery Team. 91 p. plus Appendices.

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.

Paulsen, C.M., and T.R. Fisher. 2001. Statistical relationship between parr-to-smolt survival of Snake River spring–summer Chinook salmon and indices of land use. Transactions of the American Fisheries Society 130: 347-358.

Pess, G. R., P. M. Kiffney, M. C. Liermann, T. R. Bennett, J. H. Anderson, and T. P. Quinn. 2011. The influences of body size, habitat quality, and competition on the movement and survival of juvenile coho salmon, Oncorhynchus kisutch, during the early stages of stream recolonization. Transactions of the American Fisheries Society 140:883–897.

Peterman, R.M. 1981. Form of random variation in salmon smolt-to-adult relations and its influence on production estimates. Canadian Journal of Fisheries and Aquatic Sciences 38:1113-1119.

Petrosky, C.E., and H.A. Schaller. 2010. Influence of river conditions during seaward migration and ocean conditions on survival rates of Snake River Chinook salmon and steelhead. Ecology of Freshwater Fish 19(4): 520–536.

Petrosky, C.E., H.A. Schaller and P. Budy. 2001. Productivity and survival rate trends in the freshwater spawning and rearing stage of Snake River Chinook salmon (*Oncorhynchus tshawytscha*). Canadian Journal of Fisheries and Aquatic Sciences 58:1196-1207.

Roni, P., T. Bennett, R. Holland, G. Pess, K. Hanson, R. Moses, M. McHenry, W. Ehinger, and J. Walter. 2012. Factors affecting migration timing, growth, and survival of juvenile coho salmon in two coastal Washington watersheds. Transactions of the American Fisheries Society 141: 890-906.

Rothman, K.J. 1990. No adjustments are needed for multiple comparisons. Epidemiology 1:43-46.

Schaller, H.A., and C.E. Petrosky. 2007. Assessing hydrosystem influence on delayed mortality of Snake River stream-type Chinook salmon. North American Journal of Fisheries Management 27:810–824

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.

Scheuerell, M.D., Zabel, R.W. & Sandford, B.P. 2009. Relating juvenile migration timing and survival to adulthood in two species of threatened Pacific salmon (Oncorhynchus spp.). Journal of Applied Ecology 46: 983–990.

Skalski, J. R, S. G. Smith, R N. Iwamoto, J. G. Williams, and A. Hoffmann. 1998. Use of passive integrated transponder tags to estimate survival of migrant juvenile salmonids in the Snake and Columbia rivers. Can. J. Fish. Aquat. Sci. 55: 1484-1493.

Smith, S.G., W.D. Muir, J.G. Williams, and J.R. Skalski. 2002. Factors associated with travel time and survival of migrant yearling Chinook salmon and steelhead in the lower Snake River. North American Journal of Fisheries Management 22: 385-405.

Smith, S.G., W.D. Muir, E.E. Hockersmith, R.W. Zabel, R.J. Graves, C.V. Ross, W.P. Connor, B.D. Arnsberg. 2003. Influence of river conditions on survival and travel time of Snake River subyearling fall Chinook salmon. North American Journal of Fisheries Management 23: 939-961.

Tuomikoski, J., J. McCann, B. Chockley, H. Schaller, P. Wilson, S. Haeseker, J. Fryer, C. Petrosky, E. Tinus, T. Dalton, R. Ehlke and R. Lessard. 2012. Comparative Survival Study (CSS) of PIT-tagged Spring/Summer Chinook and Summer Steelhead. 2012 annual report. BPA Contract # 19960200. Prepared by Comparative Survival Study Oversight Committee and Fish Passage Center. 164 pp. plus appendices. (http://fpc.org/)

UCSRB (Upper Columbia Salmon Recovery Board). 2014. Integrated Recovery Program Habitat Report. Available at http://www.ucsrb.org/Assets/Documents/Library/Reports/2014 Habitat Report - DMG.pdf