Appendix G

Comments and response from ISRP/ISAB

FISH PASSAGE CENTER

2501 SW First Avenue, Suite 230, Portland, OR 97201-4752
Phone: (503) 230-4099 Fax: (503) 230-7559
http://www.fpc.org
e-mail us at fpcstaff@fpc.org

August 23, 2002

Northwest Power Planning Council
Attention Judi Hertz
Response to ISRP
851 SW 6th Avenue, Suite 1100
Portland, Oregon 97204

RE: Project ID: 199602000 – Comparative Survival Study (CSS) of Hatchery PIT tagged chinook and the Comparative Survival Study Oversight Committee.

Dear Ms. Hertz:

Attached, please find the response to ISRP comments on the subject proposal.

Sincerely,

Michele DeHart

Response to ISRP comments

Project ID 199602000

Comparative Survival Rate Study (CSS) of PIT tagged Chinook & Comparative survival Study Oversight Committee

1. ISRP Comment: "The response must include an outside peer review of the estimation process by a qualified statistician(s) or there must be a programmatic review by the ISRP allowing adequate time for careful evaluation of the estimation process before a positive recommendation for funding can be given. Previous reviews by the ISAB and the ISRP resulted in the conclusion that the overall design of the data collection was adequate to meet the primary objectives of the project, but that the statistical properties of the proposed analysis procedures (mathematical formulas) should be further investigated before conclusions are based on data from this study. The previous ISRP and ISAB reviews did not approve the specific mathematical formulas in the reports issued by this project. Adequate review of the proposed analysis procedures is not feasible in the time allocated for the review for all proposals in the Mainstem and System wide Province."

Response: The study has been reviewed in detail by the ISAB on January 14, 1997, and January 8, 1998, and most recently in December 2001. John Skalski, University of Washington, provided the most recent review comments on the present study design, on December 3, 2001. A copy of those comments and the response to comments are attached. In addition, those comments and the response to those comments were appended to the annual report for 2001, which is available at http://www.fpc.org/fpc_docs/css/CSS_Report_FINAL.pdf in Appendix F.

The CSS Oversight Committee is amenable to outside independent reviews and to the ISRP detailed review discussed in their comments. The CSS Oversight Committee is scheduled to discuss the statistical and study design details with the ISRP on September 24, 2002 to facilitate the ISRP detailed review. Additionally, in response to Question # 4 posed by the ISRP, the Oversight Committee plans to begin work to publish results this winter. A broad range of peer review of statistical analysis and methodology will occur through that process.

2. ISRP Comment: "When will the project end? The reason for the project stated on page 2 is to answer, *can transportation of fish to below Bonneville compensate for the effect of the hydro system on juvenile survival rates of Snake River spring and summer chinook salmon during their downstream migration?* It appears that the direction of the project is changing to the point that the proposal should be considered a new proposal. The project began in 1996 yet the proposal notes a rather tentative goal on page 2, and repeated on page 3, *This study is intended to begin to provide the basis for the Mainstem Monitoring and Evaluation (M&E) Program's analysis of long term alternatives for recovery of depressed listed and unlisted stocks of chinook and steelhead."*

Response: This is an ongoing, long-term project, which monitors and evaluates salmon survival (smolt to adult) related to existing hydrosystem management actions (in-river migration and transportation) across a broad range of environmental conditions (e.g., runoff volumes, estuary/ocean). The project has maintained a consistent scope, which has since its inception included the identified transportation question but also several questions which are outlined in tasks and objectives of the proposal (see proposal Section 9 f). These include upstream-downstream comparisons, the development of long-term, consistent, time series of SARs, and the hydro system passage history of smolts. The CSS Oversight Committee previously responded to this question of project duration by the Northwest Power Planning Council (September 8, 1997 memo, DeHart to Casavant) as follows: "*The Salmon Managers initially proposed the PIT tagging at hatcheries as a means of evaluating mitigation measures aimed at recovery of listed wild chinook. Since recovery will take many years, there will be the need for the release of marked fish for the evaluation of recovery measures. Therefore, we will consider this study a long-term effort. Although hatchery stocks are predominately used now, as wild stock population sizes increase, they would be considered for tagging. The key element of this PIT tagging effort is to provide a level of consistent marking over time to address the effects of the primary mitigation measures. This long-term study is designed to conform with and compliment the NPPC adaptive management approach as outlined in the draft framework paper*." The ISAB review (January 8, 1998) also recommended a long-term, expanded CSS project (recommendation 2): "*So long as the present configuration and operation of the hydroelectric system exists, extend (or continue) PIT tagging to include naturally reproducing populations of spring chinook whenever population sizes may permit. Continue PIT tagging other life history types, and extend PIT tagging to other life history types of other species of salmon, including steelhead, whenever possible*."

The direction of the project is essentially the same as proposed in 1996 and 1997; however, the project has proposed additions of specific study populations to better meet the project goals, respond to project reviews by the ISAB and other reviewers, and adapt to changes in the Fish and Wildlife Program, additional ESA listings and regional programs. The key response variables have continued to be empirical smolt-to-adult return rates (SARs) compared to those needed for survival and recovery, and SAR comparisons between transport and inriver migration routes and upstream and downstream populations. The project has contained since its inception the task of exploring feasibility of developing lower river wild spring chinook index stocks to estimate smolt-to-adult return rates to compare with those of Snake River wild stocks. The current proposal, which adds steelhead groups, is consistent with the original project vision and the specific recommendation of the ISAB cited above.

The initial and present intent of this study is "*to begin to provide the basis for the Mainstem Monitoring and Evaluation (M&E) Program's analysis of long term alternatives for recovery of depressed listed and unlisted stocks of chinook and steelhead*." The basic challenge identified by the ISRP is that some components of a mainstem / systemwide M&E program are in place (including the CSS study), but the overall M&E program is not. Clearly, these component programs (including CSS) will need to mesh functionally in the future for a successful systemwide M&E program. As discussed below, formally combining projects does not seem to be necessary or beneficial at this stage so long as data collection and analytical activities are closely coordinated through the proposed umbrella project.

3. ISRP Comment: "The response should contain a careful self-review evaluating the advantages and disadvantages of combining this project with the CBFWA proposal #35033 to form a system wide monitoring and evaluation project."

Response: The CBFWA proposal #35033 for collaborative, systemwide monitoring and evaluation (if funded) would provide a framework within which the CSS (and other projects of similar scale) could operate to monitor and evaluate life cycle survival of listed and unlisted Columbia Basin salmon, steelhead (as well as resident species). Note that the CBFWA proposal did not propose to incorporate administration and implementation of projects like CSS, but rather to integrate Tier 1, 2 and 3 data from these component projects into a systemwide M&E program, and make recommendations for filling critical information gaps related to key management questions facing the region.

Until a systemwide M&E program is actually established, there does not seem to be any advantage to combining the ongoing CSS project with an un-funded proposal such as #35033. In the future, an advantage of combining this project with the CBFWA proposal #35033 might be to ensure project coordination and to prioritize CSS M&E activities. The alternative model is to keep projects separate but have close coordination between the CBFWA M&E project and the various components (including CSS) to ensure efficiency of data collection and analyses. The disadvantage to combining CSS with CBFWA proposal #35033 is primarily one of logistics of project administration and implementation. The scale of CSS is currently workable, with implementation carried out by the Smolt Monitoring Program, and project design, data analyses and oversight carried out by an interagency oversight committee. We foresee no advantages to CSS project administration or implementation from a formal incorporation of CSS into the CBFWA project, because the existing logistical burden would simply fall to the CBFWA project (and subsequently back to the Smolt Monitoring Project). Potential benefits to the CSS study design or data analyses tasks from combining projects could be achieved alternatively through coordination between the CSS project and the CBFWA proposed M&E project, especially considering the overlap of sponsoring agencies and biologists/biometricians on the two projects.

4. ISRP Comment: "The proponents should summarize progress toward publication of the results and methods in the peer reviewed literature, if any attempt has been made."

Response: A part of the CSS results concerning survival rates by route of passage has been published in the North American Journal of Fisheries Management (Budy et al. 2002). However, the majority of the methods and results are contained in the report "Comparative Survival Study of PIT tagged spring/summer Chinook Status Report for Migration Years 1997-2000 Mark/Recapture Activities" in great detail (Bouwes et al. 2002). The CSS oversight committee has been planning to submit a couple of publications, one on the methodologies and another on the results of basinwide comparisons for spring/summer chinook survival rate patterns. The publications rely on finishing the analysis of the non-parametric bootstrap technique for confidence limits for smolt-to adult return rates. In addition, we could not publish results in previous years because the adult returns were not complete until 3 years after marking. Therefore, in order to have three years of data the returns were not complete until 2002. We anticipate submitting these manuscripts for publication this winter.

5. ISRP Comment: "It was mentioned that bootstrapping would be used to obtain confidence intervals on the point estimates and we agree that this may be an appropriate procedure. However, the problem is deeper than estimation of variances. The formulas proposed are ratios of ratios and the magnitude of mathematical bias in the point estimates should also be evaluated. In addition, maximum likelihood estimators and perhaps others

should be developed and contrasted to the proposed ad hoc estimators to determine the most accurate and precise estimates possible with the available data."

Response: The ISRP agrees that the bootstrap may be an appropriate procedure for estimation of variance, but they would like to see an evaluation of potential bias in SARs, ratios of SARs, and the delayed mortality index D. The CSS researchers realize that there is a potential for biases in the estimation process that should be evaluated. For example, estimating the number of smolts in the T_0 (total transported in LGR equivalents) and C_1 (in-river migrating smolts detected at a transportation site in LGR equivalents) categories requires unbiased estimates of survival from Lower Granite Dam tailrace to Lower Monumental Dam tailrace (this expands to McNary Dam tailrace in years that springtime transportation at McNary occurs). As part of the estimation process, we look for patterns in the survival estimates between these dams that may be reflective of potential biases. An unbiased estimate of the number of smolts in the C_0 (in-river migrating smolts not detected at a transportation site in LGR equivalents) category requires unbiased survival estimates to produce results in LGR equivalents and an unbiased estimate of the population of PIT tagged fish at Lower Granite Dam (undetected and detected fish). Most of the variance and potential bias of the estimated number of smolts in Category C_0 will arise from the estimation of population at Lower Granite Dam.

We ran simulations of the process of estimating the number of undetected wild fish at Lower Granite Dam, which included seasonally and randomly varying detection probabilities, smolt travel times, and survival rates. The results suggest that our proposed method results in very small (< 1%) bias in estimates of undetected smolts at Lower Granite, with 95% confidence intervals well within ± 10% of the true value. This method must be used for wild fish, and can also be used with hatchery fish.

The ISRP recommends that we should develop maximum likelihood estimators and contrast them to our "ad hoc" estimators to determine which provides more accurate and precise parameter estimates. However, some of the quantities we already estimate, such as reach survival rates, in fact use maximum likelihood estimation, and the Lower Granite Dam population estimates are generated using components that are maximum likelihood estimators (*e.g.,* estimated collection efficiency). It is these estimates that determine the accuracy and precision of the estimated smolt numbers. These estimates in combination with the actual count data create the estimated number of smolts in each category. This is not an "ad hoc" approach as implied by the ISRP, but rather a set of computational formula based on the underlying probabilities of survival between dams, probability of collection at a dam, and probability of being transported once collected at a dam.

Where practicable, theoretical formulas for variance and/or profile confidence intervals from maximum likelihood estimation (MLE) will be employed with the original data to compare with estimates of variance and confidence intervals generated from the bootstrap program. Likelihood profiles for SARs (where the denominator is known with little error) can be generated using the binomial probability distribution and observed releases and recaptures. Variance for log-transformed ratios of SARs with denominators that are presumed to be known with little error [*e.g.,* $SAR(T_{LGR})$ and $SAR(C_1)$] can be estimated with the formula derived from the ratio of two binomial random variables [see Equation (1) of Townsend and Skalski (1997)]. Additionally, MLE for ratios of these SARs will be performed using a likelihood formula similar to Equation (14) of Townsend and Skalski (1997), generating likelihood curves and support functions, which will give means and confidence intervals which can be compared to those generated from the bootstrap. If the bootstrap estimates of these relatively simple SAR and T/C estimates exhibit low bias and robust confidence intervals, it will provide assurance that more involved estimation procedures (*e.g.*, for D) are reasonable.

Because estimates of in-river survival from Lower Granite Dam tailrace to Bonneville Dam tailrace (LGRBON reach) have generally required some extrapolation of survival across sections of river for which no direct estimate is possible, there is the potential for biases to enter into the estimation of D. In years prior to 1998, there were greater chances of biases in these expansions because of the limited PIT tag detection capabilities at John Day and Bonneville dams, compared to 1998 and subsequent years. In 1998 and subsequent years the distance of river over which in-river survival has had to be extrapolated has been reduced, thus reducing the potential for biases in the LGRBON reach survival estimate. In the bootstrapping program, we have added a feature that allows the researcher to pre-select the number of reaches over which to use existing estimates of in-river survival and to choose among alternative methods of extrapolation. This will allow us to compare the sensitivity of the resulting LGRBON reach survival estimate to the amount of reach (distance) being extrapolated, and the method used.

6. ISRP Comment: "Why is NMFS not on the interagency Comparative Survival (CSS) Oversight Committee? It seems that they are one of the primary users of the results and should be directly involved in oversight of the project."

Response: NMFS was invited to join the Oversight Committee at the inception of the Committee and the CSS study. NMFS declined to participate in day-to-day Oversight Committee discussions. However, NMFS Science Center staff participated in the early stages of study statistical design development. NMFS has not been excluded from the Oversight Committee and has a standing invitation to join if they so desire. NMFS as well as any other agency or individual is provided the opportunity to review and comment on the CSS, annual report, annual proposal study designs and any other aspect of the CSS. NMFS has taken the opportunity to provide comments on this study through the NMFS ESA Section 10 permit process for the CSS.

7. Action Agencies/ NMFS RME Group Comments: "The RME Hydro subgroup recognizes that the proposed research has the potential to provide data and estimates useful in satisfying elements in those RPAs, Hydro-related RME RPAs 185, 187, 188, and 189. The smolt survival estimates have further application in the context of testing compliance with the Hydro performance standards as noted for other proposals in this review. The proposal was thorough in specifying sample sizes comprising key index treatment groups. However it would be beneficial if that information was translated into precision estimates. Alternatively power analyses for key hypothesis tests could be presented to demonstrate the estimates will be satisfactory for evaluating key hypothesis remaining in the region. This would also aid in assessing the utility of the information in performance tests that would be performed at the check-ins."

Response: The CSS provides data useful to addressing hydro-related RPA 185 (SARs of in-river and transported smolts and associated estimation of delayed mortality of transported fish), RPA 187 (relation between ocean entry timing and SARs of in-river and transported smolts), RPA 188 (SARs of lower Columbia River basin wild stocks for use in evaluating effects of hydro system on upriver stocks), and RPA 189 (SARs of smolts with different passage histories through the hydro system, including effects such as number of bypasses detected and which particular bypasses detected). Through the large scale PIT tagging of hatchery yearling chinook and steelhead, the CSS will provide a database containing smolt passage histories and adult return histories. For Snake River basin smolts, this database will provide direct comparisons of SARs of in-river and transported smolts with a 90% power of detecting differences of at least 50% between the two outmigration routes as long as the smaller SAR does not drop below 1%. For Mid-Columbia River basin smolts, this database will provided direct comparisons of SARs of in-river smolts against the COE's McNary Dam transported smolts with a 90% power of detecting differences of at least 30% between the two outmigration routes as long as the smaller SAR does not drop below 1%. Once any other specific hypothesis of interest to the region is formulated, it would be feasible to evaluate the power of testing that hypothesis using the CSS database. However, we cannot guarantee that the power will be as high for those specific tests if the numbers of smolts available for these new hypothesis tests are much lower than the number of smolts required for the original hypotheses. The PIT tagging of wild smolts at tributary traps will provide marked fish in addition to those NMFS is PIT tagging at the dams for use in estimating SARs from and back to Lower Granite Dam. From the composite of wild stocks, estimates of SARs and ratios of SARs will be possible, but given the uncertainty of collecting large enough numbers of fish of wild origin, the power of the tests will typically be lower than what is possible with the fish of hatchery origin.

The precision of the estimated SARs for in-river and transported smolts will be obtained through bootstrapping techniques. The bootstrap will also provide precision of the ratios of SARs and the associated delayed mortality "D" index. The bootstrap can be an effective tool to obtain a valid measure of variability in a parameter, even when that parameter is a computation based on a set of values, each of which must be estimated. For example, when the ratio of returning adults to a known (fixed) number of smolts is used to generate an estimated SAR, the underlying binomial distribution may be used to obtain the associated measure of precision of the SAR estimate. However, when the number of smolts must also be estimated, the underlying distribution of the ratio of two estimated parameters becomes more complex. For these situations, the non-parametric bootstrap technique is useful (Dixon 1993). Likewise, the ratio of pairs of these SARs (*e.g.*, ratio of transported LGR-LGR SAR to in-river LGR-LGR SAR) would form a complex underlying distribution for which the use of the bootstrap is a preferred approach. This is also true of the estimation of delayed transportation mortality, the D parameter or the ratio of BON-LGR SARs . Programmers at the Fish Passage Center are currently writing a computer program to perform bootstrapped estimates of variance and confidence intervals for individual SARs, ratios of SARs, and D. The next CSS annual status report will contain bootstrapped estimates of precision for all parameters presented. This will allow NMFS to assess the utility of using the CSS's estimated parameters at their periodic check-ins.

References

Bouwes, N., C. Petrosky, H. Schaller, P. Wilson, E. Weber, S. Scott, and R. Boyce. February 2002. Comparative Survival Study of PIT tagged spring/summer Chinook Status Report for Migration Years 1997-2000Mark/Recapture Activities. BPA contract number 8712702. Columbia Basin Fish and Wildlife Authority, Fish Passage Center Report. www.FPC.ORG.

Budy, P., G. P. Thiede, N. Bouwes, C. E. Petrosky, and H. Schaller. 2002. Evidence linking delayed mortality of Snake River Salmon to their earlier hydrosystem experience. North American Journal of Fisheries Management Vol. 22 (1), pp 35-51.

Dixon, P. M. 1993. The bootstrap and jackknife: describing the precision of ecological indices. Pages 290-318 in S. M. Scheiner and J. Gurevitch, editors. Designs and analysis of ecological experiments. Chapman and Hall, New York

Townsend, R.L. and J. L. Skalski. 1997. A comparison of statistical methods of estimating treatment-control ratios (transportation benefit ratios) based on spring chinook salmon on the Columbia River, 1986-1988. The design and analysis of salmonid tagging studies in the Columbia basin, Volume IX. BPA Project Number 91-051, Portland OR.

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
ISAB@nwcouncil.org

ISAB Review of the 2005 Comparative Survival Studies' Annual Report and Applicability of Comparative Survival Studies' Analysis Results

Robert Bilby
Susan Hanna
Nancy Huntly, Chair
Stuart Hurlbert
Roland Lamberson
Colin Levings
William Pearcy
Thomas Poe
Peter Smouse

Charles Coutant, Ad Hoc Member
Richard Alldredge, Ad Hoc Member, ISRP

ISAB 2006-3
March 15, 2006

ISAB Review: The 2005 CSS Annual Report and Applicability of CSS Analysis Results

ISAB Review: The 2005 CSS Annual Report and Applicability of CSS Analysis Results

Executive Summary

On December 20, 2005, the Council requested that the Independent Scientific Advisory Board (ISAB) review the 2005 Annual Report for the Comparative Survival Study (CSS) prepared by the Fish Passage Center (FPC) and the Comparative Survival Study Oversight Committee, as well as critical comments on the draft of that report by the Bonneville Power Administration (BPA) and NOAA Fisheries. The CSS is a field study, begun in 1996, that addresses important and technically complex issues regarding the survival of PIT-tagged Spring/Summer Chinook and PIT-tagged Summer Steelhead through the Columbia River hydrosystem from juveniles through returning adults. The study focuses on relative survival of fish that traveled downstream as juveniles by alternative routes (e.g., in river, transported, different routes of dam passage, and different numbers of dams passed). The results can have important implications for operation of the hydrosystem to ensure protection and propagation of anadromous salmonids. The Council expressed a desire to aid resolution of disputes over the study by obtaining the ISAB review.

The Council asked that the ISAB assess the overall integrity and scientific soundness of the CSS report and address the following specific questions:

> *1. Are the design, implementation, and interpretation of the statistical analyses underpinning the report based on the best available methods? Does the ISAB have suggestions for improving the analyses?*
>
> *2. What is the applicability of the CSS results, taking into account whatever scientific criticisms of the analyses that the ISAB decides are valid, if any? In other words, what weight should the analyses be given and what qualifiers should be considered when using the analyses for decision-making?*

The ISAB accepted the assignment on January 12, 2006 and received a briefing on the CSS Annual Report from the study's Principal Investigators on January 27th. The ISAB considers that there are two parts to this review: (1) review of the 2005 CSS Annual Report and (2) a determination of the utility of the CSS comparative survival estimates for various management and hydrosystem operational decisions.

The ISAB finds that the CSS is an ambitious, long-term study that is being criticized because its objectives are not yet fully met, despite prodigious efforts in both the field and in complex data analyses. The CSS has used the PIT-tag technology to mark and track individual salmon and steelhead through their smolt-to-adult life stages. Expectations of this mark-recapture technology exceed the results that are practically attainable, and its use is still evolving. The CSS study participants have been major players in this evolution. We find the present annual report to be a further incremental step in the direction of documenting different survival rates of different stocks under different migration conditions. That the present report is not a perfect reconstruction of differential survival histories is largely a result of the current analytical capabilities and available sample sizes. The deficiencies seem to be highlighted in some aspects because of experimental design and analytical approaches taken by the authors. The ISRP comment from their 2002 review still applies that "the formulas [used to compute relative survival rates] are complicated, convoluted, and in general, very unsatisfactory from a statistical point of view."

Specific Responses to the Council's Questions

1. Are the design, implementation, and interpretation of the statistical analyses underpinning the report based on the best available methods? Does the ISAB have suggestions for improving the analyses?

All in all, the design, implementation, and interpretation of the *statistical analyses* underpinning the report are very good. Nonetheless, there are broader concerns over the design of the study such as sample size, sampling sites, time periods for analyses, and other features. Improvements can be made, and our recommendations follow.

Since the region is unwilling to conduct the manipulative experiments in the hydrosystem that the ISAB and ISRP have recommended for many years, the CSS is doing the next best thing. That is, the study is following as many fish through their life cycle as possible, calculating the survival, and comparing outcomes.

2. What is the applicability of the CSS results, taking into account whatever scientific criticisms of the analyses that the ISAB decides are valid, if any? In other words, what weight should the analyses be given and what qualifiers should be considered when using the analyses for decision-making?

The ISAB believes the Council should view the CSS as a good, long-term monitoring program, the results of which should be viewed with increasing confidence as years pass. Under scrutiny from periodic peer reviews and agency comments, the methods should improve and the results become ever more valuable. The project is definitely worthy of Council support.

The Council's question is difficult to answer with the present annual progress report. The project needs a synthesis report that clearly describes the analytical methods and summarizes the project results in a holistic way for its decade of effort.

The ISAB recognizes a disconnect between the present status of results and much of the decision-making that takes place regarding hydrosystem operations and fish protection. Although the project is making good progress at addressing such issues as the value of transportation and the relative survival from different passage routes, many relationships between survival and specific operational alternatives or environmental features during migration cannot be resolved when data are aggregated simply by year of migration. For this information to be most useful for making management decisions, aggregations of data within years and across years for different operational options and environmental constraints should be pursued. We encourage the project to move in that direction.

The results of the CSS appear to indicate that PIT-tagged fish do not have the same survival rate as untagged fish. This conclusion is not emphasized by the current progress report, but it has major implications for many uses of the PIT-tag technology. Comparisons among PIT-tagged groups of fish are probably appropriate, but extrapolations of the results from PIT-tagged fish to untagged populations should be made with caution.

Recommendations

- It has been ten years since the CSS was initiated. The report the ISAB reviewed was the latest in a series of annual progress reports, and thus lacking a holistic perspective. The ISAB recommends that the CSS produce a ten-year summary report providing an in-depth description of methods and detailed analyses and interpretation of the data in a retrospective style.

- The CSS needs to more effectively present the methodologies used in their analyses so the criticism of complicated and convoluted formulas can be avoided. The scattered explanations in several annual progress reports could be consolidated in the ten-year summary recommended above.

- The ISAB agrees with critics who express concern that two downriver sites (Carson Hatchery and John Day River) are probably insufficient to give accurate upriver-downriver comparisons of SARs. This concern is bolstered by the variability among upriver hatcheries shown by the CSS data. For this upriver-downriver comparison to be generally accepted, it seems prudent to add more downriver sites in the future.

- Data on size of all PIT-tagged fish from hatcheries and other release sites should be included in the report in much greater detail. Size at release may be a significant factor in differential SARs. The ISAB recommends including a specific section in the report focusing on the potential effects of size at release on survival of all PIT-tagged fish.

- Aggregation of data solely by juvenile migration year should be supplemented with analyses that group data on environmental and operational factors that may be amenable to control.

- Assumptions inherent in the analyses should be specifically tested, with continued vigilance toward avoiding bias.

- Pre-assigning the intended routes of passage at the time of release into inriver and transport groups would greatly simplify calculation of SARs and eliminate much criticism of current methods that are unnecessarily complex. This modification to the study design is scheduled for implementation in 2007, but should begin in 2006, if feasible.

- Analyses could emphasize more diverse metrics of differential survival, thus avoiding the criticism that the project staff focuses mainly on contentious issues such as the relative survival of transported and in-river migrants (T/C ratios) and differential delayed mortality between transported and in-river migrants (*D*). Passage routes, numbers of dams bypassed, distance from ocean, different hatchery practices, and other features have been explored beyond the issue of transportation.

- The CSS should be supplemented by funded research into analytical methods that can improve, and hopefully simplify, the mathematical and statistical approaches currently in use. It is not clear from available information whether the problem is that the formulas are unnecessarily complicated, inappropriately specified, or just not well explained (see bullet #2 above).

- More attention should be given by the CSS and the region as a whole to the apparent documentation that PIT-tagged fish do not survive as well as untagged fish. This point has major implications for all uses of PIT-tagged fish as surrogates for untagged fish.

I. Introduction and Background

Review Assignment

On December 20, 2005, the Council requested that the Independent Scientific Advisory Board (ISAB) review the 2005 Annual Report for the Comparative Survival Study (CSS) prepared by the Fish Passage Center (FPC) and the Comparative Survival Study Oversight Committee. The CSS is a field study of the survival of PIT-tagged Spring/Summer Chinook and PIT-tagged Summer Steelhead through the hydrosystem from juveniles through returning adults, with a focus on relative survival of fish that traveled as juveniles by alternative routes (e.g., in river, transported, different routes of dam passage, and different numbers of dams passed). The annual report reviews recent mark/recapture activities and bootstrap analysis for generating confidence intervals.

The CSS is important, as it is one of the few organized attempts to systematically release PIT-tagged, hatchery-reared fish, and wild smolts into the Columbia River for the purpose of monitoring and evaluation. Most aspects of the study, from its design and methods to the analytical results, have been strongly debated in the Region because the relative survival rates of salmonids under different hydrosystem operations and environmental constraints is at the heart of water and fish management policies.

In response to the release of the draft version of this annual progress report, both the Bonneville Power Administration and NOAA Fisheries provided the FPC with letters setting forth both broad concerns and detailed criticisms of the findings and results reported in the draft report. Before finalizing the report, the FPC provided detailed responses to both Bonneville and NOAA Fisheries addressing their concerns. The Council expressed its wish to contribute to the resolution of these important and technically complex issues by having the ISAB conduct its own review of the final progress report and the attendant letters. In conducting the review, the Council asked that the ISAB assess the overall integrity and scientific soundness of the CSS report and address the following specific questions.

1. Are the design, implementation, and interpretation of the statistical analyses underpinning the report based on the best available methods? Does the ISAB have suggestions for improving the analyses?

2. What is the applicability of the CSS results, taking into account whatever scientific criticisms of the analyses that the ISAB decides are valid, if any? In other words, what weight should the analyses be given and what qualifiers should be considered when using the analyses for decision-making?

The ISAB accepted this important assignment on January 12, 2006 and received a briefing on the CSS Annual Report from the study's Principal Investigators on January 27th. The ISAB considers that there are two parts to this review: (1) review of the 2005 CSS Annual Report and (2) a determination of the utility of the CSS comparative survival estimates for various management and hydrosystem operational decisions.

The CSS was initiated in 1996 by the Northwest fishery agencies and tribes as a long-term study to estimate survival rates over different life stages of spring and summer Chinook salmon produced in hatcheries in the Snake River basin and selected lower hatcheries in the lower Columbia River. The study has expanded somewhat to encompass wild Chinook salmon and steelhead, and the mix of hatcheries has changed with experience. The premise of the research was that, through use of PIT tags implanted in juveniles at the point of release from hatcheries or rearing facilities, the survival of unique groups of fish could be determined as they passed through PIT-tag detectors in juvenile bypasses at dams or in adult fish ladders on their return. From these survival rates it was hypothesized that one could quantify differential survival according to passage route. Of particular interest were differences in survival related to distance from the ocean, between transported and in-river fish and the delayed effects of hydrosystem passage (by juveniles) on adult returns.

Previous Reviews

Both the ISAB and the ISRP previously reviewed the CSS study proposals in 1998 (ISAB 1998) and 2002 (ISRP 2002) and the recommendations from those reviews were generally as follows (recommendations are provided in full in Appendix A):

In 1998, the ISAB supported funding of the study. They recommended including naturally reproducing populations as well as hatchery fish and suggested that other life-history types of Chinook salmon and steelhead be included. They recommended quantifying survival from tributary hatcheries to Lower Granite Dam and McNary Dam, and through the entire hydrosystem when sufficient detectors were functional. They encouraged attempts to compare survival of PIT-tagged fish to untagged fish or fish tagged by other methods. The ISAB also saw this as a way to coordinate the PIT-tagging efforts of many agencies and to provide an opportunity for periodic workshops to review results.

The ISRP reviewed the continuation proposal in 2002 and also recommended funding. The "best" formulas for calculating smolt-to-adult survival rates from then-available data were judged "complicated, convoluted, and in general, very unsatisfactory from a statistical point of view." It was noted that arguments over these methods would likely continue and spawn even more detailed arguments and counter-arguments. Much of the difficulty lies in small sample sizes due to both numbers of fish tagged and the number of detections. Improved detection at Bonneville Dam was recommended. The ISRP recommended more research on mathematical and statistical methods both within this project and outside it for estimating life-cycle survival.

II. Review of 2005 CSS Annual Report

Methods (Chapter 2)

There are three principal issues over the study's methods. One concerns the selection of hatcheries (or other release sites), especially for comparisons between smolts with long passage routes through the hydrosystem and those migrating from lower in the basin with few dams to pass. Another relates to the mathematical and statistical methods employed in the analyses, including potential biases and the types of aggregation of data for summaries. A major point raised by NOAA Fisheries is the unreliability of the PIT-tag method to represent the survival of untagged fish (the CSS data indicate that PIT-tagged fish do not survive as well as untagged fish, and therefore are not adequate surrogates for untagged fish in the population).

Some study methods are not fully described in this annual progress report. We did not seek out previous annual progress reports to fill in the information gaps. This difficulty begs for a summary report that can provide a more complete description of methods.

It would be useful to have the SARs analyzed as a function of size at release. This could be tested for rather than just presenting size data. Also, data on size of all PIT-tagged fish from hatcheries and other release sites should be included in much greater detail than median lengths at tagging reported in Table 2 (e.g., include mean lengths, weights, and ranges). Sizes at release may be a significant factor in differential SARs from various sources. Fish size is generally not accorded much significance in the CSS studies despite a well-known survival advantage for larger fish. As raised in comments by NMFS, these size effects need to be given more consideration in further analyses. The ISAB recommends including a specific analyses focusing on the effects of size at release on SAR values of all PIT-tagged fish.

The numbers of fish available for tagging is a major constraint. As tables 2-5 demonstrate, the number of tagged fish vary considerably by location and year. The study participants have had to be opportunistic despite an intended experimental design. To their credit, they appear to have been quite successful in obtaining numerous stocks and years to compare.

Holdovers (fish not migrating fully through the hydrosystem in the year of initial outmigration; Connor et al. 2002) cause methodological problems. The authors have tried to account for these fish in different ways in this and the previous annual report. They believe the present method has less bias for estimating survival. This needs to be evaluated in later years.

We admire the study participants for attempting to segregate fish among their several migration-route histories. Although the term "destined" seems too strongly pre-ordained for the current methods of release and tracking, fish do have the three options listed: in-river by non-bypass routes, in-river through dam bypasses, or routed to transportation at the collector dams. They have these options at most dams (not all dams have facilities to collect fish for transportation), thus expanding the number of possible migration histories. Equipment failures, changes in protocols at a particular dam from year to year, and other irregularities complicate matters even more. This is a real “haystack” of PIT-tag data from which to extract the key “needles” in the form of meaningful comparisons of survival among both source groups and passage histories.

As in the comments by BPA and NMFS, we are critical of the authors’ choice to summarize SAR results only on an annual basis. The determinants of SAR likely vary as much with the environment within a migration year as between years, and these could be tested. The environmental status and hydrosystem-operating mode at the specific time a fish migrates through the system represents the features that are most relevant to survival and are specific targets for modification, rather than average conditions over a migration year. It has been an ongoing criticism of the FPC that they do not further refine their data analyses to within-year conditions (e.g., the ISAB’s comments on the FPC flow augmentation analyses reported in ISAB flow augmentation reviews (ISAB 2004-2)).

We recognize the problems presented by segregating migration histories within years. For example, fish from a release batch disperse in the river and do not all pass a dam at the same time, and therefore individuals experience different environmental and operational histories. However, further breakdown by operational modes or environmental features (such as temperature ranges) could greatly enhance the value of further analyses of the CSS data. The annual summaries can be considered as broad “first cuts” that may be modified by these additional analyses.

The evolving nature of these analyses is reflected in Table 8, which shows older and more recent estimates of the comparison of the differential delayed mortality between transported and in-river fish (*D*). Despite the number of significant figures reported, the overall number can change, as the influences on it are better understood and included in calculations. Although labeled as a “correction” based on comments on the draft report we see the change as progressive improvement (they may change again).

The study has necessarily aggregated batches of tagged fish, as described at the bottom of page 12. The authors seem to have accounted for this in a reasonable way.

As an overall perspective, there is no way of avoiding the realization that there are a lot of assumptions inherent in the study, from tagging through analyses and presentation of data. Further research should test these assumptions, or tag a sufficient number of appropriate fish so that empirical data can replace assumptions.

Much of the continuing controversy is related to the mathematical and statistical methods employed. We agree with the earlier ISAB comment that the "formulas are complicated, convoluted, and in general, very unsatisfactory from a statistical point of view." That said, we think the FPC response to the issues raised by NMFS and BPA is quite good. Where questions of bias in estimators are raised, the primary issue appears to be estimating SAR starting from the population at Lower Granite Dam rather than from other projects. However, the ISAB found the explanation by the CSS scientists as to why the estimate was made in this manner to be reasonable.

There are assumptions made no matter which method is proposed for estimation. For example, the CSS makes the assumption that the transportation proportion for the unmarked population of each hatchery group and the aggregate wild group is approximately the same. Also, it is assumed that the PIT tagged and untagged smolts have the same probability of surviving to and being collected at the dams in the hydro system. These assumptions should be tested.

With respect to the assertion that the PIT tagging reduces survival (see NOAA Fisheries' comments below), we are concerned about the basic premise of the CSS, namely that PIT-tagged fish can serve as surrogates for the unmarked population. If this assertion stands up to further scrutiny, then use of PIT tags should be restricted to comparisons among PIT-tagged groups, and not with unmarked fish.

The use of the bootstrap method to estimate confidence intervals is appropriate. The methodology is now widely used in many statistical applications.

The ISAB hopes the sponsors will more effectively present the methodologies used in the next (2006) Annual Report or in the 10-year summary report we recommend so the criticism of complicated and convoluted formulas can be avoided.

Results (Chapter 3)

The level of scientific satisfaction with the results varies among the species and stocks analyzed. In some cases the results as presented are fairly robust; in other cases where data are scant, trends may be visible but lack statistical significance. The authors present what they have.

Wild Chinook

The problem of small sample sizes for wild Chinook is clearly illustrated by Table 9, which presents the age composition of their PIT-tagged returns. Although a few years had three-digit numbers per age category (1999, 2000, 2002), other years had single- or double-digit numbers. Expansions, while logical, still do not avoid the problem of having few adult returns. Regrettably, it is the wild Chinook that suffer most severely from this concern.

The low return rates of tagged wild Chinook cause the SAR estimates to be very uncertain. The 90% confidence limits of the transport SAR calculations (Table 11) show very wide ranges. What reasonable conclusions can one make when the 90% confidence ranges from zero to over 3? The results do more to demonstrate the *lack* of ability to determine the true SAR than anything. The authors recognize this difficulty in the text on page 15, and we can take their analyses as a straightforward presentation of the SAR values they calculated using limited data.

The authors were criticized for comparing their calculated SAR values (inexact as they probably are) to the 2% for stable stocks and 4% for recovery recommended by Marmorek et al. (1998). We find no fault with their flagging their calculated values near 1% as a likely problem. We agree with critics of the study that there are better estimates now of stock-specific returns needed for stable populations and recovery, and better calculations of SAR values would be an improvement. But the general trend is unsettling and the CSS results should be taken in their intended context.

The consistent trend in the comparison of SAR values for smolts collected at a collector dam (C_1) and those not detected (C_0) (page 16) also is troubling, despite understood problems with the data. A difference of 25% might just be real. (The table referred to should be Table 12, not Table 10).

In our view, the scant data provide essentially no meaningful information on the relative survival of transported smolts and in-river migrants (T/C ratio) for wild Chinook salmon in all years except 2001 (Figure 4). That year most smolts were transported because of extremely low river flows and high temperatures for in-river migrants, and the transport SAR was high. The values of the differential delayed mortality between transported and in-river migrants (*D*) have a similar limitation, as the authors note.

We are inclined to view the further analysis of wild Chinook data on pages 19-24 as not warranted based on the scant amount of data available. Perhaps we do not follow the intent of the authors in this section. Further combining of SARs, T/Cs, and *D*s to come up with sample sizes suitable for statistical analysis seems to us to be inappropriate. The more fruitful direction for the longer term would seem to be to tag more fish in order to match these values with specific operational and environmental regimes that could (at least for operations) be modified to obtain better survival.

Hatchery Chinook

The foundation of data for hatchery Chinook salmon is much better than for wild Chinook (Table 17). However, when taken to the level of specific source hatchery (Table 19), in many cases the data look nearly as sparse as for wild Chinook.

We did not specifically critique the authors' results or discussion of each specific hatchery. The variation among hatcheries is rather expected, based on different rearing conditions, fish size at release, distance from the ocean, etc. The authors seem to have made logical attempts to explain differences in SAR performances. It is interesting that the Rapid River Hatchery seems to be the closest surrogate for wild Chinook. Size effects noted earlier probably deserve more attention.

The T/C ratios among hatcheries are nearly all above 1, indicating superior survival of the transported fish. The ratios are not far above 1, however, and only the estimated error bounds get above 2 (the expected T/C in the absence of *D*).

Wild Steelhead

The numbers of returning adult steelhead are even fewer than for wild Chinook, and thus the results are even less reliable. We view these results as merely presentation of what is available, rather than providing a strong case for any conclusion. Within the limitations of the data, some of the same trends appear as for Chinook, such as higher SAR values for fish not detected as smolts, somewhat higher SARs for transported fish (for steelhead this was above 2 three of 5 years, excluding 2001), and widely varying D values. The issue of residualism is important for steelhead, as the authors point out.

Hatchery Steelhead

Low numbers of fish make this analysis problematic. Small sample sizes yield no statistically significant results. However, the authors carry through with the same analyses as for the other groups. The most interesting suggestion is that a possible relationship between fish detected at collector dams and those undetected through the hydrosystem appears to have disappeared in 2000 and 2002.

Adult Drop-out Rates (Chapter 4)

The potential for loss of adults migrating upstream being influenced by the outmigration experiences of the fish as smolts has been raised in the region. We were pleased to see the adult PIT-tag detection data used to track adult upstream movements and losses. The data seem to support conclusions that dropout is higher where there is a fishery (not unexpected), hatchery fish dropped out somewhat more than wild (not stressed by the authors), and that transported fish had a somewhat higher dropout rate than in-river fish. The comparisons in this report just scratch the surface of what can be learned from these data. More important than the Transport/In-river comparisons are potential insights into migration rates at different flows and other environmental differences. Perhaps the emphasis on "survival" in the CSS led to the more narrow focus.

Hatchery-to-Hatchery SARs for Various Hatcheries (Chapter 5)

A basic premise of the CSS was that different survival rates could be calculated for each hatchery from which smolts were released. After many adjustments for terminal fisheries and other factors, this chapter seems to be a straightforward presentation of the SAR values from hatchery back to hatchery for five hatcheries. The problem of small sample sizes is evident. In order to have enough fish for hatchery comparisons, the authors did not do a transported vs. in-river comparison.

Upriver-Downriver Comparisons (Chapter 6)

A prime motivation for the CSS was the hypothesis that the SARs for salmonids that must pass downstream through the hydropower system as juveniles would be lower than those for fish passing no or few dams. To test this hypothesis, there must be adequate representation from both upriver and downriver fish sources.

We concur with critics who express concern that the two downriver sites (Carson Hatchery and John Day River) are probably too few to give accurate upriver-downriver comparisons. This concern is bolstered by the variability among upriver hatcheries shown by the CSS data. For this upriver-downriver comparison to be reliable, it seems prudent to add more downriver sites in the future.

Partition of results into common-year effects and differential mortality as carried out by Deriso et al. (2001) and this study appears reasonable and justified, despite criticisms from Williams et al. (2005). As an editorial note, “fig.y” and later “fig yy” need their numbers.

Estimates of differential upriver-downriver mortality based on spawner-recruit and PIT-tag SAR values provide useful confirmation during the one year of overlap (2000). It would be useful to continue these parallel analyses. We do not understand, however, how averaging 1.48, 0.78, and 1.18 supports the conclusion that upriver stocks survive “about 1/3 as well as John Day populations for these years.”

We were puzzled that the conclusions listed for this chapter did not mention the upriver-downriver comparison for which the chapter was titled. Instead, the conclusions relate to common survival patterns estimated by the two techniques, comparison of wild and hatchery fish, and high correlations among populations. It would have been informative and appropriate to include the comparative survival information (upriver populations survived about 1/3 as well) in the conclusions.

Simulated PIT-tag data to test CJS survival estimates (Chapter 7)

In principle, one can test the reliability of analytical methods by developing simulated data sets and conducting analyses on them. We generally concur that testing the analytical approach with simulated data should provide a useful evaluation of the approach. The present section provides insufficient information, however, to understand what is being done. The abbreviation CJS needs to be defined.

ISAB Evaluation of Comments by BPA and NOAA Fisheries

BPA Comments

BPA was critical of the observational nature of the CSS, the use of a "heuristic analytical approach" devoid of a statistical model, bias in the estimates that lead to incorrect conclusions, misguided emphasis on *D*, a misguided upriver-downriver comparison, and generally flawed and skewed interpretations that minimize the benefits of transportation and the return rates of salmonids. It provided its own mathematical derivation of transported SAR as an appendix.

BPA's initial criticism that the CSS cannot make direct causal inferences about any particular natural or anthropogenic factor is technically correct, as is the need for manipulative and replicated experiments in order to do so. However, the ISAB and its precursor advisory bodies have requested such manipulative and replicated experiments in the FCRPS for more than a decade, and the requests have been refused by BPA and other action agencies as impractical. BPA is criticizing the CSS for deficiencies in their study when these deficiencies have been caused largely by BPA policy decisions. What the CSS is doing is consistent with its initial study proposal, continuing objectives, and periodic technical reviews.

We do not fault the CSS for its empirical approach. First, the CSS authors do not merely compare hatchery-to-hatchery SAR values, but try several measures of survival along the migration corridor. Survival to Lower Granite Dam is used as a more reliable measure than returns to the hatchery of origin, for example. The CSS has standardized much of its data to the LGR site. We do not see that the approaches used in the CSS analysis are appropriately characterized as biased. As the BPA commenter notes, the issue is somewhat moot because the CSS results do show advantages for transportation in some years, especially in the drought year of 2001.

We do not see that the CSS has focused on *D* as a primary gauge of the effectiveness of transportation. It seems to be presented as one measure along with others. We believe that use of multiple metrics benefits the comparisons. In addition, delayed mortality is real. Therefore, why shouldn't one calculate the difference in this delayed mortality between transported and in-river fish? We note that the CSS has updated its estimates of *D* based on comments, which we take as a sign of continual improvement.

Some inconsistency between earlier progress reports and this one are to be expected. That's why they are "progress reports." This criticism is one reason why the ISAB sees the need for a ten-year summary report as well as the incremental annual reports.

We concur that the upriver-downriver comparison has problems. The BPA commenter correctly criticizes the CSS for relying on just one downstream hatchery when the upstream hatcheries showed such wide variation in results. But the BPA comment does not acknowledge that the CSS also used the John Day River stock for the downriver set. The Hilborn et al. (1993) paper cited by BPA (without reference) does not eliminate the possibility that information other than that used by Hilborn et al. could show differences between upriver and downriver performance. We would encourage the CSS participants to build on this critique and bolster the downriver samples.

NOAA Fisheries Comments

The NOAA Fisheries comments reflected their belief that the analyses in the progress report are incomplete, do not fully support the findings in the executive summary and chapters, and lack a holistic approach to analyzing all available data. They argue for more in-depth analyses and broader discussion of all relevant data on the effects of the hydropower system on salmonid stocks. They opine that PIT-tagged fish do not represent the untagged populations, that the CSS made selective use of data, that statistical significance is used inconsistently, and that there are biases in the comparisons between treatments and controls. A major point is that the PIT-tagged fish really do not provide a true representation of the untagged population, based on the CSS data. In addition to these general topics, they provided detailed comments by section.

The ISAB suggests that the NOAA Fisheries' expectation that the present annual progress report be a holistic evaluation of all data is unrealistic. That criticism would be more appropriate for a final or periodic summary report. An annual progress report is, by design, of more limited scope. We do agree, however, that a holistic summary is sorely needed after 10 years of work and incremental progress reports.

The NOAA commenter states that the PIT-tagged fish do not represent the survival of the untagged population, while the CSS premise is that they would and the report implies that they do. This is an important difference. In the NOAA Fisheries' comments (and in the technical memo they cite), they note that the PIT-tagged fish returned at about ½ the rate of untagged fish. The data to make these comparisons is in the CSS report, but the CSS authors do not make the comparisons. We agree with NOAA Fisheries that this difference is not trivial and that the CSS must discuss it as well as simply present results. In our view, however, the CSS quite fairly presents the PIT-tag data as its best estimate, although admittedly imperfect. The difficulty comes from comparing the results to the published 2% value for sustainability of a population (tagged and untagged).

We concur that there is some vagueness in statements about statistical significance. On some points, the CSS report simply relies on overlap of the 90% confidence limits. In other places it is not so clear. The CSS could improve this aspect of its reporting. Statistical significance should be tested for and the nature and level of significance of the tests reported.

We concur that size of fish matters and that more attention should be placed on fish sizes in subsequent CSS analyses.

We agree that the Executive Summary could better reflect the results of Chapter 3 in regard to the degree to which hatchery fish can be used as surrogates for wild fish. Nonetheless, the statement that the CSS continues to evaluate this seems appropriate.

As NOAA Fisheries comments, the bullets for Chapter 3 could better represent the text. But these bullets need to be understood as brief summaries of what the text reports.

As we noted before, we concur that use of only one hatchery for the downriver comparison is not good practice, considering the variation seen in results for upstream hatcheries.

The detailed comments are valuable for the CSS to consider as it moves along with the work.

III. ISAB Answers to Council's Questions

1. Are the design, implementation, and interpretation of the statistical analyses underpinning the report based on the best available methods? Does the ISAB have suggestions for improving the analyses?

All in all, the design, implementation, and interpretation of the *statistical analyses* underpinning the report are very good. Nonetheless, there are broader concerns over the design of the study such as sample size, sampling sites, time periods for analyses, and other features. Improvements can be made, and our recommendations follow.

Since the region is unwilling to conduct the manipulative experiments in the hydrosystem that the ISAB and ISRP have recommended for many years, the CSS is doing the next best thing. That is, the study is following as many fish through their life cycle as possible, calculating the survival, and comparing outcomes.

The study design could be improved in several ways. Adding more downriver hatcheries to make more valid upstream/downstream survival comparisons. Much more attention should be given to the size of tagged fish at various release locations, because survival is known to be affected strongly by fish size. The data could be aggregated to more closely meet the needs of hydrosystem managers. Whether by design or implementation, the aggregation of data simply by year of outmigration is insufficient to resolve many of the important issues related to environmental influences and hydrosystem operations. The numbers of fish tagged may never be sufficient for resolving in-season patterns of survival. However, as data are accumulated over more years, it may be feasible to partition analyses into environmental or operational categories across years to obtain more functional correlations. Having a controlled and manipulated experimental design would be preferable (as BPA asserts), but the chance of this happening is slim. Repeated entreaties by the ISAB, its predecessor advisory bodies and the ISRP have all been met with objections to the effect that such a system wide experiment is not possible to manage (although we note that the region managed to implement high spill in 2005 on court order, although no planned experiments were conducted). The opportunistic approach of documenting survival under whatever conditions are dealt seems to be the only alternative.

Implementation would be improved by tagging more fish (particularly wild), but there is likely a limit to the amount that can be accomplished due to manpower limitations. The study managers have been quite opportunistic in arranging tagging and in coordinating tagging efforts among many different entities. Pre-assignment of fish to either inriver or transport passage routes at the time of release would greatly improve study design and make the analyses and results more transparent. Assignment of passage route at release is planned for implementation in 2007 (i.e., a given tag number would really be “destined” to be shunted to a particular route, if possible). This modification should be implemented in 2006, if possible.

The data analyses require extensive statistical manipulations to extract useful information from the mass of PIT-tag detections. We can only agree with the earlier ISRP comment that the "formulas are complicated, convoluted, and in general, very unsatisfactory from a statistical point of view." Pre-assignment of fish to inriver and transport groups at time of release should help. The study participants have gone to great lengths to seek ways to analyze the data appropriately. Bootstrapping confidence limits is a major improvement. We do not find any particular bias in the analyses or interpretations. Likewise, we see no inherent problem with the assumptions, and some assumptions will always have to be made. These assumptions should be tested as the project progresses.

Taken alone, the current progress report does not adequately present the analytical methods and some data presentations are difficult to follow (e.g., labeling axes as log survival instead of actual survival). The ISAB encourages the sponsors to more effectively present the methodologies in a summary report (perhaps as part of the 2006 Annual Report) so the methods of analysis can be better understood.

2. What is the applicability of the CSS results, taking into account whatever scientific criticisms of the analyses that the ISAB decides are valid, if any? In other words, what weight should the analyses be given and what qualifiers should be considered when using the analyses for decision-making?

The Council’s question is difficult to answer with just the present annual progress report. The value of this project for informing management decisions on the hydropower system would be greatly enhanced if a synthesis report were produced that clearly describes the analytical methods and summarizes the project results in a holistic way for its decade of effort. We recognize that this is what NOAA Fisheries hoped to see.

The CSS is providing long-term monitoring of lifetime survival of salmon and steelhead stocks using a technology that the region has spent a great deal of money developing and implementing. As an ongoing effort, subject to periodic review and comment, it is providing an evolving picture. It would be wrong to believe that the results as of today are the end-all for making decisions about the operation of the hydrosystem. The CSS is learning as it goes, which is to be expected. More years and more analyses of specific questions are needed.

Because the CSS is focusing on annual data, the relationships to specific operational and environmental factors within years are not addressed. As commenters have pointed out, these more specific correlations would be more useful for guiding operational decisions. The ISAB recognizes a disconnect between the present status of results and much of the decision-making that takes place regarding hydrosystem operations and fish protection. Although the project is making good progress at addressing such issues as the value of transportation and the relative survival from different passage routes, many relationships between survival and operational or environmental features during migration cannot be resolved when data are aggregated simply by year of migration. For this information to be most useful for making decisions, aggregations of data within years or across years for different operational options and environmental conditions need to be pursued. Even after aggregating the available, relevant data across several years, there may not be a sufficient number of tag detections to make such correlations for all important combinations of operational status and environmental conditions. Either more fish need to be tagged or correlations made after more years of data for which operational and environmental modes can be grouped. The former would be the more expeditious approach.

IV. ISAB Conclusions and Recommendations

The CSS is an ambitious, long-term study that is being criticized because its objectives are not yet fully met, despite prodigious efforts in both the field and in complex data analyses. It has used the PIT-tag technology to mark and track individual salmon and steelhead through their smolt-to-adult life stages. Expectations of this mark-recapture technology exceed the results that are practically attainable, and its use is still evolving. The CSS study participants have been major players in this evolution. We find the present annual report to be a further incremental step in the direction of documenting different survival rates of different stocks under different migration conditions. That the present report is not a perfect reconstruction of differential survival histories is largely a result of the current analytical capabilities and available sample sizes. The deficiencies seem to be highlighted in some aspects because of experimental design and analytical approaches taken by the authors. The ISRP comment from their 2002 review still applies that “the formulas are complicated, convoluted, and in general, very unsatisfactory from a statistical point of view.”

The Council should view the CSS as a good, long-term monitoring program the results of which will become increasingly valuable to managers as years pass. Scrutiny from periodic peer reviews and agency comments will help ensure that the methods and analytical approaches improve. The project is definitely worthy of Council support.

Recommendations

- It has been ten years since the CSS was initiated. The report the ISAB reviewed was the latest in a series of annual progress reports, and thus lacking a holistic perspective. The ISAB recommends that the CSS produce a ten-year summary report providing an in-depth description of methods and detailed analyses and interpretation of the data in a retrospective style.

- The CSS needs to more effectively present the methodologies used in their analyses so the criticism of complicated and convoluted formulas can be avoided. The scattered explanations in several annual progress reports could be consolidated in the ten-year summary recommended above.

- The ISAB agrees with critics who express concern that two downriver sites (Carson Hatchery and John Day River) are probably insufficient to give accurate upriver-downriver comparisons of SARs. This concern is bolstered by the variability among upriver hatcheries shown by the CSS data. For this upriver-downriver comparison to be generally accepted, it seems prudent to add more downriver sites in the future.

- Data on size of all PIT-tagged fish from hatcheries and other release sites should be included in the report in much greater detail. Size at release may be a significant factor in differential SARs. The ISAB recommends including a specific section in the report focusing on the potential effects of size at release on survival of all PIT-tagged fish.

- Aggregation of data solely by juvenile migration year should be supplemented with analyses that group data on environmental and operational factors that may be amenable to control.

- Assumptions inherent in the analyses should be specifically tested, with continued vigilance toward avoiding bias.

- Pre-assigning the intended routes of passage at the time of release into in-river and transport groups would greatly simplify calculation of SARs and eliminate much criticism of current methods that are unnecessarily complex. This modification to the study design is scheduled for implementation in 2007, but should begin in 2006, if feasible.

- Analyses could emphasize more diverse metrics of differential survival, thus avoiding the criticism that the project staff focuses mainly on contentious issues such as the relative survival of transported and in-river migrants (T/C ratios) and differential delayed mortality between transported and in-river migrants (*D*). Passage routes, numbers of dams bypassed, distance from ocean, different hatchery practices, and other features have been explored beyond the issue of transportation.

- The CSS should be supplemented by funded research into analytical methods that can improve, and hopefully simplify, the mathematical and statistical approaches currently in use. It is not clear from available information whether the problem is that the formulas are unnecessarily complicated, inappropriately specified, or just not well explained (see bullet #2 above).

- More attention should be given by the CSS and the Region as a whole to the apparent documentation that PIT-tagged fish do not survive as well as untagged fish. This point has major implications for all uses of PIT-tagged fish as surrogates for untagged fish.

References

ISAB 1998. Review of Comparative Survival Rate Study of Hatchery PIT Tagged Chinook. Report ISAB 98-1, Northwest Power Planning Council and National Marine Fisheries Service, Portland, Oregon.

ISRP 2002. Final Review of Fiscal Year 2003 Mainstem and Systemwide Proposals. Report ISRP 2002-14. Northwest Power and Conservation Council, Portland, Oregon.

Appendix A: Previous Review Comments by ISAB and ISRP

ISAB Comments (ISAB 1998)

- Fund the proposed study.

- So long as the present configuration and operation of the federal hydroelectric system exists, extend (or continue) PIT tagging to include naturally reproducing populations of spring chinook whenever population sizes may permit. Continue PIT tagging other chinook life history types, and extend PIT tagging to other life history types of other species of salmon, including steelhead, whenever possible.

- Apply enough PIT tags to spring chinook production from Kooskia, Pahsimeroi, McCall, Sawtooth, and Clearwater (Powell, Crooked River and Red River Ponds) hatcheries to estimate survival to Lower Granite Dam. Whenever possible apply enough PIT tags to spring chinook at these hatcheries to estimate survivals to McNary Dam.

- Compare rates of return to each hatchery of PIT tagged and untagged adults to establish degree of comparability of survivals of PIT tagged juvenile salmon to survivals of juveniles not PIT tagged. To investigate rate of shedding of PIT tags through the adult stage, and where straying of adults from another hatchery is possible, investigate thermal mass marking of all hatchery production. Where smolt to adult survival of PIT tagged fish is compared to that of coded wire tagged (CWT) fish, develop a procedure to study tag loss and to compare rate of return of PIT to CWT within the hatchery release.

- Make estimates of survival applicable to the entire Snake-Columbia River federal hydroelectric system as soon as possible.

- Promote coordination and cooperation among agencies applying PIT tags and other marks by including a list of other agencies marking salmon and steelhead of the same origin in the proposal, along with comments from those other agencies. Sponsor an interagency workshop on the use of tagging data at five-year intervals. The workshop would produce consensus recommendations and procedures for coordinating tagging activities.

ISRP Comments (ISRP 2002)

Various scientists in the region, in particular scientists from the Comparative Survival Study project and NMFS, have considered the problems in estimating the LGD to LGD smolt-to-adult survival rates (SARs) from currently available data and have apparently arrived at what they consider to be the "best" formulas. Unfortunately, the formulas are complicated, convoluted, and in general, very unsatisfactory from a statistical point of view. Accordingly, there is high probability that these methods will continue to spawn arguments and counter-arguments over trivial issues that will occupy the resources of the region, because the stakes are high (e.g., high costs of spill, high costs of transportation, unknown long term effects of the non-normative transportation, high costs of flow augmentation, etc).

The long-term solutions to the mathematical and statistical problems in estimation of smolt-to-adult return rates (Bonneville to Bonneville and Bonneville to Low Granite SARs) appear to be: 1) detection of sufficient numbers of PIT tagged juveniles passing Bonneville Dam Powerhouse II at the planned corner collector; 2) estimates of mortality of fish passing via that route; 3) and/or sufficiently large sample sizes of PIT tagged fish downstream of Bonneville. The ISRP recommends that these sampling efforts for PIT tagged juveniles be given high priority by the Council and the Corps of Engineers. In particular, Task 2 of NMFS proposal #198331900 for development of PIT tag detection in the corner collector at Bonneville Dam Powerhouse II should be given high priority.

We do not provide unqualified endorsement of the particular estimation formulas that are proposed, and we recommend that continuing statistical methods research be directed at investigating the performance of various proposed estimators and possible alternatives, including but not limited to the proposed methods and planned bootstrapping. Such research on mathematical and statistical methods could be pursued by the sponsors of this project, and by others. As an aid to clarity in comparison among possible alternative analyses, we recommend that the FPC make available a single reference data set which includes all the necessary interpretation of route of passage of PIT tagged fish and culls any suspect or ambiguous data that might be subject to further interpretation. The budget for the recommended mathematical and statistical analyses is relatively minor compared to the total cost of the project so investigation of our unresolved questions about statistical methods should not require substantial reallocation of the budget in this project to ensure compatibility of objectives, common methods and protocols. This coordination could be accomplished under the favorably reviewed CBFWA proposal #35033.

w:\em\ww\isab projects and reports\1 isab final reports\isab 2006-3 css review.doc

United States Department of the Interior

FISH AND WILDLIFE SERVICE

Columbia River Fisheries Program Office
1211 SE Cardinal Court, Suite 100
Vancouver, Washington 98683

May 31, 2007

Patty O'Toole
Program Implementation Manager
Northwest Power and Conservation Council
851 SW 6th Avenue, Suite 1100
Portland, OR 97204-1348

Dear Patty,

Below is our response to the Independent Scientific Review Panel's (ISRP) review of the Comparative Survival Study (Project 19960200 – PIT tagging spring/summer Chinook). This project was recommended for funding by the Mainstem/Systemwide Review Team (MSRT) as a Core Project. It has been recommended by the MSRT to fund project 199602000 at FY 2007 level of $1,365,000.

Please let me know if you need any additional information.

Sincerely,

Howard Schaller, Ph.D.
Project Leader
US Fish and Wildlife Service
Columbia River Fisheries Program Office
1211 S.E. Cardinal Court, Suite 100
Vancouver, WA 98683
Phone:(360)604-2500
Fax: (360) 604-2505
Email:Howard_Schaller@fws.gov
http://www.fws.gov/columbiariver/

cc: Eric Merrill, NPCC
Tom Iverson, CBFWA

RESPONSE TO QUESTIONS FROM ISRP REVIEW OF PROJECT 199602000 (PIT TAGGING SPRING/SUMMER CHINOOK- Comparative Survival Study) PROPOSAL FOR 2007 TO 2009

Proposal sponsored by USFWS - Columbia River Fisheries Program Office.

In the ISRP review of the Comparative Survival Study (Project 19960200 – PIT tagging spring/summer Chinook), they stated "*this is a supportable proposal but a response is needed to address issues raised in the ISAB's recent report: Review of the 2005 Comparative Survival Studies' (CSS) Annual Report and Applicability of Comparative Survival Studies' Analysis Results* (www.nwcouncil.org/library/isab/isab2006-3.htm)."

The ISRP lists recommendations from the ISAB report to which the USFWS proposal sponsors need to make a written response before final decision is made on the funding status for this proposed study. Each of the recommendations (shown in italics) is followed by our response (normal type).

Recommendation 1:

It has been ten years since the CSS was initiated. The report that the ISAB reviewed was the latest in a series of annual progress reports, and thus lacking a holistic perspective. The ISAB recommends that the CSS produce a ten-year summary report providing an in-depth description of methods and detailed analyses and interpretation of the data in a retrospective style.

Response 1:

The CSS will produce a ten-year summary report in FY 2007, which will look in depth at issues such as fish size effects on inriver collection efficiency and subsequent SARs, seasonal trends in SARs of transported and bypassed fish, and environment's (flow, spill, and temperature) effects on in-river survival and SARs of in-river migrating smolts including both bypassed and non-bypassed fish. In addition, the computer program developed over the past two years to create simulated datasets will be used to evaluate assumptions of the Cormack-Jolly-Seber release/recapture model, and robustness of inriver survival estimates to violations of key assumptions.

Recommendation 2:

The CSS needs to more effectively present the methodologies used in their analyses (in this proposal as well as their annual report), so the criticism of complicated and convoluted formulas can be avoided. The scattered explanations in several annual progress reports could be consolidated in the ten-year summary recommended above.

Response 2:

One of the deliverables to BPA in 2006 will be a new design and analysis report that will present the methodologies in a more succinct mathematical framework. The WDFW member of the CSS Oversight Committee is working on the preparation of this document showing the likelihood function derivations of the SARs for each study category in the CSS including $SAR_1(T_0)$, $SAR_2(T_0)$, $SAR(C_0)$, and $SAR(C_1)$, plus the mathematical derivation of the formulas that estimate number of smolts in each study category, T/C ratios and D.

Recommendation 3:

The ISAB agrees with critics who express concern that two downriver sites (Carson Hatchery and John Day River) are probably insufficient to give accurate upriver-downriver comparisons of SARs. This concern is bolstered by the variability among upriver hatcheries shown by the CSS data. For this upriver-downriver comparison to be generally accepted, it seems prudent to add more downriver sites in the future.

Response 3:

Another downriver site in the Warms Springs River is planned for wild Chinook tagging for 2007 to complement the ongoing tagging in the John Day River. If additional downstream site are to be added to the CSS, then more funding must be made available. To date the CSS has not been able to fund any more tagging than has occurred since 2001.

Recommendation 4:

Data on size of all PIT-tagged fish from hatcheries and other release sites should be included in the report in much greater detail. Size at release may be a significant factor in differential SARs. The ISAB recommends including a specific section in the report focusing on the potential effects of size at release on survival of all PIT-tagged fish.

Response 4:

Based on findings published by NOAA Fisheries researchers on potential size effects on collection efficiency and subsequent survival, the CSS plans to include a chapter in the 2007 CSS Summary Report to look at the effects of size at tagging. Lengths were taken on 10% of hatchery Chinook being PIT-tagged at Dworshak, Rapid River, and McCall hatcheries during the spring tagging season. Wild Chinook that were PIT-tagged in the spring primarily at the lower tributary traps on the Salmon, Imnaha, Grande Ronde, and Clearwater rivers may be good candidates for investigation of potential effects due to size at tagging for wild Chinook stocks. Lengths of wild fish tagged during late summer to fall of the year prior to springtime migration would not reflect lengths at migration and these fish may be less useful for examining effects of length on collection efficiency and subsequent survival.

Recommendation 5:

Assumptions inherent in the analyses should be specifically tested, with continued vigilance toward avoiding bias.

Response 5:

We plan to create sets of simulated data to evaluate how sensitive CJS survival estimates are to violations of assumptions used in the estimation process. . These evaluations will be reported in the ten year CSS summary Report.

Recommendation 6:

Pre-assigning the intended routes of passage at the time of release into in-river and transport groups would greatly simplify calculation of SARs and eliminate much criticism of current methods that are unnecessarily complex. This modification to the study design is scheduled for implementation in 2007 (according to the 2005 Annual Report but this change in protocol should be indicated in the proposal).

Response 6:

Beginning with the 2006 migration year, the CSS already adopted the approach of pre-assigning a group of PIT-tagged fish to represent the untagged populations' experience through the hydrosystem and a second group of PIT-tagged fish to provide the required in-river survival estimates with the CJS release/recapture methods. Pre-assigned groups were used in the CSS for 2006 including each individual Chinook hatchery, the aggregate wild Chinook, aggregate wild steelhead, and aggregate hatchery steelhead. Two-thirds of the PIT-tags were pre-assigned to groups reflecting the untagged populations and the remaining one-third were pre-assigned to the group used to obtain inriver survival estimates. This approach will continue to be implemented in future years as well.

Recommendation 7:

Analyses could emphasize more diverse metrics of differential survival, thus avoiding the criticism that the project staff focuses mainly on contentious issues such as the relative survival of transported and in-river migrants (T/C ratios) and differential delayed mortality between transported and in-river migrants (D). Passage routes, numbers of dams bypassed, distance from ocean, different hatchery practices, and other features have been explored beyond the issue of transportation.

Response 7:

In preparing the 2007 CSS Summary Report, a 10-year synthesis of what has been learned to date from this study, we plan to explore additional metrics of differential survival, as recommended by the ISAB. In 2006, transportation began later at the Snake River collector dams, and we plan to evaluate the earlier years data with regard to whether higher overall SARs would have occurred on collected fish if all fish were bypassed until later in April before beginning transportation. These evaluations will address the question raised by the COE regarding "what to do with the collected fish – transport or bypass them?" PIT-tagged fish have been monitored at the Rapid River Hatchery outfall since 1999 and since fish volitionally exit that facilitie's pond, we plan to evaluate temporal differences in survival rates to Lower Granite and subsequent SARs for earlier, middle, and later outmigrating smolts. Smolts in study category C_0 pass the three collector dams on the Snake River inriver through non-bypass routes, either through spill or the turbines.

We plan to look at relations between estimated SAR for C_0 fish and levels of spill (volume or proportion of discharge) occurring at these dams. The question raised by NOAA Fisheries researchers that smaller fish may be prone to higher collection in the bypass, but lower overall survival will also be investigated. For wild Chinook, we will use PIT-tagged fish released from Smolt Monitoring Program traps on the lower Salmon, Imnaha, Grande Ronde, and Clearwater rivers. These fish are PIT-tagged in the spring with lengths taken on each tagged fish, and migrate to Lower Granite Dam relatively quickly so any further growth would be negligible. For hatchery Chinook, we will use PIT-tagged fish released from Dworshak, Rapid River, and McCall hatcheries. These fish are PIT tagged one to two months before release with lengths taken on 10% of the tagged fish. Some additional growth may occur between tagging and when these fish arrive at Lower Granite Dam, but it is unlikely the size differences would diminish by the time they enter hydrosystem, thus allowing a greater opportunity to see differences in collection efficiency and subsequent SARs, if they do indeed occur.

We also plan to investigate SARs (BON-BON) based on arrival timing to Bonneville Dam between C0, C1 and T0 groups of Snake River and downriver wild and hatchery Chinook.

Recommendation 8: *In addition to the ISRP recommendations, the ISAB noted that more attention should be given by the CSS and the Region as a whole to the apparent documentation that PIT-tagged fish do not survive as well as untagged fish. This point has major implications for all uses of PIT-tagged fish as surrogates for untagged fish.*

Response 8: We plan to compare SARs estimated from PIT tagged spring/summer Chinook groups with SARs estimated from untagged fish that rely upon methods outlined in Petrosky et al. (2001) and Williams et al. (2005).

<u>Other comments -- A:</u>
A timeline with years (1996 - current) should be included within the background section to improve the proposal. Details in this section are sparse and references are lacking. The proponents either assume that the reviewers know all the background and justification for this project or decided not to go through the work needed to provide the details.

<u>Response A:</u>

The project began in 1996 and has had extensive regional review. The ISAB reviewed the CSS on January 14, 1997, and followed that review with a face-to-face meeting in Spokane WA on March 10, 1997. As a result of the 1997 reviews, the ISAB was better informed on purposes of upstream/downstream portion of study. They recommended an oversight committee for the study and recommended that NMFS be represented, but attempts by CSS to include NMFS failed due to disagreements in validity of upstream/downstream comparisons. Based on the ISAB 1997 review, the CSS was consolidated from two separate BPA project numbers (#198712700 and #199602000) into one project number #199602000.

Another review by the ISAB occurred on January 6, 1998. In that review the ISAB recommended adding other species of salmon including steelhead, but to date CSS has not been able to get BPA funding for steelhead. We are attempting to add steelhead to the CSS again in the 2007 – 2009 proposal. In the 1998 review, the ISAB also concurred with shift from proportional tagging to PIT tagging a minimum of 45,000 hatchery Chinook at key study hatcheries for assessing hatchery-specific SARs. In addition, the ISAB recommended resampling or other methods for variances of SAR; thereafter CSS began work on a non-parametric bootstrap approach, which is now incorporated in CSS annual reports.

On July 16, 2002, CSS Oversight Committee members made a presentation on the estimation formulas used in the CSS plus the bootstrap used for estimating confidence interval during an ISRP review meeting. The ISRP was also briefed on the importance of T/C ratios and D in assessing management actions. The presentation was followed up with written responses by CSS to ISRP comments on August 23, 2002. Based on ISRP recommendations, the CSS Oversight Committee added a chapter to the 2002 Annual Report comparing the bootstrap with likelihood-based confidence intervals. In addition, we began programming to implement the ISRP recommendation for *Monte Carlo* simulations to assess validity of bootstrap confidence interval coverage. On September 18, 2002, the ISRP provided additional questions to CSS, which were addressed in face-to-face meeting in Seattle on September 24, 2002.

On January 27, 2006, Oversight Committee members, Tom Berggren, FPC, Howard Schaller, USFWS, Charlie Petrosky, IDFG and Paul Wilson, USFWS had a face-to-face meeting with the ISAB in Seattle, Washington. At the meeting, the Oversight Committee members delivered a presentation covering the 2005 CSS Annual Report and goals of the CSS. The Oversight Committee members answered questions about possible bias identified in the BPA/NOAA comments and asked again at the meeting by Steve Waste of the NPCC. The primary criticism from BPA/NOAA was that the estimates produced by the CSS were biased due to the estimation of the transport and inriver SARs. The Oversight Committee explained that the CSS technique appropriately answers a specific set of questions. These questions are (1) what is the SAR of fish arriving Lower Granite Dam “destined” for transportation and (2) what is the SAR of fish arriving Lower Grantie Dam “destined” to remain inriver and undetected at Lower Granite, Little Goose, and Lower Monumental dams. By starting at Lower Granite Dam we are comparing the transported and inriver fish over the same reach (i.e., from Lower Granite Dam as smolts to Lower Granite Dam as adults). The BPA recommendation is to start the estimation only after the fish to be transported are in the barge or truck. We told the ISAB that both approaches are unbiased, and the only difference is in where you want to start indexing the SAR for transported fish. Dr. John Skaski, in 2000 recommended using Lower Monumental Dam tailrace as the starting location for the inriver migrants in order to obtain an “unbiased” SAR. As we explained to the ISAB, if we take the BPA recommended transport SAR and divide it by Dr. Skalski’s recommended inriver SAR we would obtain lower T/C ratios than what we obtain when staring all fish at Lower Granite Dam. These differences still don’t mean that one method is biased and the other is not biased; instead they only reflect the differences in SARs that will be obtained when the starting location for indexing SAR changes. The difference is that the CSS approach measures the SARs that the run at large experienced for transport and inriver fish. In other words, the CSS approach is measuring transport and inriver SARs, T/Cs and D values for a set of conditions the fish experienced. Using the BPA recommended approach would be for a set of conditions the fish do not experience presently. The differences in approach become more of a philosophical question (Should we measure a set of condition that does not exist precisely, or should we measure the actual set of conditions that fish experience with slightly less precision?) than a statistical question.

A large proportion of the presentation was geared at informing the ISAB on the purposes and modeling approach used in the upstream/downstream comparison. We presented the ISAB with the background, hypotheses, and rationale behind the design of the CSS. The CSS is a coordinated regional effort under the auspices of a regional oversight committee and is closely tied to the goals of the Mainstem Monitoring and Evaluation Program. The ISAB asked many questions and the session ended with them having a much better understanding of the background, history, motivation for the study and evaluation techniques used in the CSS project. Thus far, ten years of juvenile marking have been completed. Adult returns from migration years 1996 to 2003 have been analyzed in five Project Status Reports completed in 2001, 2002, 2003, April 2005, and December 2005. At the recommendation of the ISAB during the project review meeting of January 26, 2006, a more detailed retrospective compilation of what has been learned in the CSS from these ten years of study will be produced in FY 2007.

Other comments -- B:

The project history section consists of only a few sentences and is lacking sufficient detail to provide project accomplishments and give adequate justification for continued support. For such a long-running project there have been a number of important accomplishments and completed documents that need to be listed in this section.

Response B:

CSS was begun in 1996 with approximately 5% of hatchery spring/summer Chinook production above Lower Granite Dam PIT-tagged in numbers proportional to total hatchery release. All fish were returned-to-river at Snake River collector dams for inriver survival estimation. In 1997 the CSS was modified to fixed release numbers at four specific hatcheries – Dworshak, Rapid River, McCall, Imnaha, and Lookingglass (onsite release and Imnaha acclimation pond). Beginning in that year the study was expanded to include the routing of a proportion of PIT-tags to transportation at the collector dams. From 1997 to 1999, Lower Granite Dam was considered the primary transportation site with the overall transportation quota met either by that site alone (1997) or that site in combination with Little Goose Dam for part of the season (1998 and 1999). By migration year 2000, it was determined that potential differences in site-specific SARs may occur among the three collector dams on the Snake River and so for all years from 2000 to 2005, an equal proportion of first-time detected PIT-tagged at Lower Granite, Little Goose, and Lower Monumental dams has been routed to transportation (proportions ranging from 50% to 67% depending on year and species/rearing type). When ODFW ceased making the Lookingglass Hatchery onsite releases in 1999, the CSS switched to the Lookingglass Hatchery release at Catherine Creek Acclimation Pond in 2001. Beginning in 2002 the CSS began coordinating with other research programs to allow a portion of their PIT-tagged wild Chinook to be routed to transportation at the Snake River collector dams, as well as fund additional PIT tagging of wild Chinook at key Smolt Monitoring Program traps and provide 14,500 PIT tags at other IDFG tributary traps to supplement ongoing tagging activities there. The

CSS began a similar effort of coordinating with other research programs to allow a portion of their PIT-tagged wild steelhead to be transported in 2003.

PIT tagging of hatchery Chinook at downstream hatchery facilities began in 1996 at Round Butte Hatchery (Deschutes River) and Cowlitz Hatchery (Cowlitz River), with Carson Hatchery (Wind River) added in 1997. The Cowlitz Hatchery tagging occurred only in 1996 and 1997, and the Round Butte Hatchery tagging occurred only in 1996, 1997, and 1998. The difficult logistics in obtaining fish to tag coupled with BKD levels at the hatchery caused us to discontinue using Round Butte Hatchery, while at Cowlitz Hatchery, the primary concern was that the spring Chinook production was more ocean type than stream type in rearing and not as directly comparable to the upstream hatchery fish as Carson Hatchery fish. The Carson Hatchery stock has been PIT tagged for the CSS in each year of study since 1997. Wild Chinook PIT tagged in the John Day River under an ODFW contract with BPA have provided a source of fish for SAR computation since 2000 in the CSS. These downstream stocks have provided SAR information that has been used in spawner/recruit modeling efforts to investigate hydrosystem effects on Chinook stocks originating in tributaries above Lower Granite Dam.

In 2006 at the request of the ISAB and NOAA representative to the ISAB, the CSS began the approach of pre-assigning PIT tags at time of tagging to one of two groups – one group reflecting the untagged population in which case any fish entering the bypass/collection system at Lower Granite, Little Goose, or Lower Monumental Dam will be transported whenever the run-at-large is being transported, and the other group will be bypassed back-to-river if entering the bypass/collection system at any of these sites. In both groups, PIT-tagged fish passing through spill or turbines at a given dam will be undetected at that site. The bypass group consisting of undetected and detected fish remaining inriver will provide the CJS inriver survival estimates between release and Lower Granite Dam tailrace and between Lower Granite Dam and Bonneville Dam for use in indexing SARs to Lower Granite Dam and computations of the delayed mortality parameter (D).

The CSS has produced five project status reports (completed in October 2000, February 2002, November 2003, April 2005, and December 2005) and a report documenting the CSS design and analysis (completed in 2001). References for these documents are listed below. Bootstrap confidence intervals for study parameters have been computed and presented in the past three project status reports. A flowchart of the simulation program was presented in Chapter 6 of the 2003/04 CSS Annual Report. A series of simulation runs to evaluate validity of T_0, C_0 and C_1 SARs estimates and proper coverage of confidence intervals resulting from bootstrap program is planned for the 2006 CSS Annual Report, with further work on this topic continuing into the proposal years of 2007 to 2009. The 2007 CSS Summary Report will provide be a more detailed retrospective compilation of what has been learned in the CSS from these ten years of study as recommended by the ISAB following the January 26, 2006, review meeting on the CSS. In addition, an updated CSS design and analysis report is being produced for 2006 showing a detailed mathematical treatment of the estimators used in the CSS for SARs, T/C ratios, and D.

The CSS Oversight Committee also conducted a workshop in February 2004 on effects of hydrosystem configuration and operation on salmon and steelhead survival. Objectives were to: synthesize results of CSS and other research studies; document and assess evidence related to various factors that can affect survival rates over different life history stages, including hydrosystem passage, delayed mortality, time of ocean entry and travel time; produce a report synthesizing and assessing the evidence for and against hypothesized mechanisms for differential survival (hatchery-wild; upstream-downstream) and SARs; and provide a foundation for a series of publications in peer-reviewed journals. Workshop proceedings were published as Marmorek et al. (2004).

Reference

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.

Williams, J.G., S.G. Smith, R.W. Zabel, W.D. Muir, M.D. Scheuerell, B.D. Sandford, D.M. Marsh, R.A. McNatt, and S. Achord. 2005. Effects of the Federal Columbia River Power System on salmonid populations. NOAA Technical Memorandum NMFS-NWFSC-63. (http://www.nwfsc.noaa.gov)

Addendum

Reference list of CSS produced documents:

Berggren, Thomas and Larry Basham – Fish Passage Center. October 2000. Comparative Survival Rate Study (CSS) of Hatchery PIT Tagged Chinook, 2000 Annual Report, Status Report for Migration Years 1996–1998 Mark/Recapture Activities. Report to Bonneville Power Administration, Contract No. 8712702, 58 pages. Available at http://www.fpc.org/

Berggren, Tom – Fish Passage Center, Nick Bouwes – Eco Logical Research, Howard Schaller, Paul Wilson – USFWS, Charlie Petrosky – IDFG, Earl Weber – CRITFC, Shane Scott – WDFW, Ron Boyce – ODFW. 2002. Comparative Survival Rate Study (CSS) 2002 Design and Analysis Report. Report to Bonneville Power Administration, Contract No. 00006203, Project No. 199602000, 34 electronic pages (BPA Report DOE/BP-00006203-3)

Bouwes, Nick – Eco Logical Research, Charlie Petrosky – IDFG, Howard Schaller, Paul Wilson – USFWS, Earl Weber – CRITFC, Shane Scott – WDFW, Ron Boyce – ODFW. February 2002. Comparative Survival Rate Study (CSS) of Hatchery PIT tagged Chinook, 2001 Annual Report, Status Report for Migration Years 1997–2000 Mark/Recapture Activities. Report to Bonneville Power Administration, Contract No. 00006203, Project No. 199602000, 100 electronic pages (BPA Report DOE/BP-00006203-2). Available at http://www.fpc.org/

Berggren, Thomas, Henry Franzoni, and Larry Basham – Fish Passage Center, Paul Wilson and Howard Schaller – USFWS, Charlie Petrosky – IDFG, Earl Weber – CRITFC, Ron Boyce – ODFW, Nick Bouwes – Eco Logical Research. November 2003.

Comparative Survival Study (CSS) of PIT Tagged Spring/Summer Chinook, 2002 Annual Report, Migration Years 1997-2000 Mark/Recapture Activities and Bootstrap Analysis. Report to Bonneville Power Administration, Contract No. 00006203, Project No. 199602000, 85 pages. Available at http://www.fpc.org/

Berggren, T., H. Franzoni, L. Basham, P. Wilson, H. Schaller, C. Petrosky, K. Ryding, E. Weber, and R. Boyce. April 2005. Comparative Survival Study (CSS) of PIT-tagged Spring/Summer Chinook. 2003/04 Annual Report, Migration Years 1997-2002 Mark/Recapture Activities and Bootstrap Analysis. BPA Contract # 19960200. Available at http://www.fpc.org/

Berggren, T., H. Franzoni, L. Basham, P. Wilson, H. Schaller, C. Petrosky, E. Weber, and R. Boyce. December 2005. Comparative Survival Study (CSS) of PIT-tagged Spring/Summer Chinook and PIT-tagged Summer Steelhead, 2005 Annual Report, Mark/Recapture Activities and Bootstrap Analysis. BPA Contract # 19960200. 107 pages. Available at http://www.fpc.org/

Marmorek, D.R., M. Porter, I.J. Parnell and C. Peters, eds. 2004. Comparative Survival Study Workshop, February 11-13, 2004: Bonneville Hot Springs Resort. Report compiled and edited by ESSA Technologies Ltd., Vancouver, B.C. for Fish Passage Center, Portland, OR and the US Fish and Wildlife Service, Vancouver, WA. 137 pp.

Example publications and reports using CSS information:

Budy, P., G.P. Thiede, N. Bouwes, C.E. Petrosky, and H. Schaller. 2002. Evidence linking delayed mortality of Snake River salmon to their earlier hydrosystem experience. North American Journal of Fisheries Management 22:35-51.

Budy, P. and H. Schaller (in review). EVALUATING THE POTENTIAL OF TRIBUTARY RESTORATION TO INCREASE THE OVERALL SURVIVAL OF SALMON. Ecological Applications

Marmorek, D.R., M. Porter, I.J. Parnell and C. Peters, eds. 2004. Comparative Survival Study Workshop, February 11-13, 2004: Bonneville Hot Springs Resort. Report compiled and edited by ESSA Technologies Ltd., Vancouver, B.C. for Fish Passage Center, Portland, OR and the US Fish and Wildlife Service, Vancouver, WA. 137 pp.

Muir, W. Marsh, B. Sandford, S. Smith and J. Williams (in press). Post-Hydropower System Delayed Mortality of Transported Snake River Stream-type Chinook Salmon: Unraveling the Mystery. Transactions of the American Fisheries Society

Paulsen, C. M., and T. R. Fisher. 2005. Do Habitat Actions Affect Juvenile Survival? An Information-Theoretic Approach Applied to Endangered Snake River Chinook Salmon. Transactions of the American Fisheries Society 134:68-85.

Peters, C.N. and D.R. Marmorek. 2001. Application of decision analysis to evaluate recovery actions for threatened Snake River spring and summer chinook salmon (*Oncorhynchus tshawytscha*). Canadian Journal of Fisheries and Aquatic Sciences 58:2431-2446.

Schaller, H.A and C.E. Petrosky *in review*. Evaluating the influence of delayed mortality on Snake River stream-type Chinook salmon (*Oncorhynchus tshawytscha*). Submitted to North American Journal of Fisheries Management

Williams, J.G., S.G. Smith, R.W. Zabel, W.D. Muir, M.D. Scheuerell, B.D. Sandford, D.M. Marsh, R.A. McNatt, and S. Achord. 2005. Effects of the Federal Columbia River Power System on salmonid populations. NOAA Technical Memorandum NMFS-NWFSC-63. (http://www.nwfsc.noaa.gov)

Wilson, P.H. 2003. Using population projection matrices to evaluate recovery strategies for Snake River spring and summer Chinook salmon. Conservation Biology 17:782-794.

Appendix H

Response to Comments

FISH PASSAGE CENTER
1827 NE 44th Ave., Suite 240, Portland, OR 97213
Phone: (503) 230-4099 Fax: (503) 230-7559
http://www.fpc.org
e-mail us at fpcstaff@fpc.org

August 31, 2007

Robert J Austin
Deputy Director of Fish and Wildlife
Bonneville Power Administration
PO Box 3621
Portland, Oregon 97208-3621

Dear Mr. Austin:

Thank you for your review of the Draft, Ten Year Retrospective Summary Report. The following response was developed by the Comparative Survival Study Oversight Committee, (Committee) comprised of, the Columbia River Inter-tribal Fish Commission, the Washington Department of Fish and Wildlife, the Oregon Department of Fish and Wildlife, the Idaho Department of Fish and Game, and the US Fish and Wildlife Service. As you are aware the Comparative Survival Study is a joint project of the agencies and tribes. The study design, the implementation of the study and the analysis are carried out collaboratively among the sponsoring fish and wildlife management agencies. The Committee has developed the following response to your general comments, followed by the response to each specific comment.

General Comments

As with past BPA review comments, we found several comments which will be helpful in improving the overall strength of the final report. However, many of the BPA general comments summarized in the cover letter are presented in such general terms without an explicit context that they are difficult to address. They are presented as sweeping conclusions of a critical nature without any basis provided. Further, some of the general statements are inaccurate and some of the reviewers' specific comments are erroneous.

Transparency, reproducibility, data, detailed methods, tagging results

A majority of the BPA conclusion comments addresses the issues of transparency, reproducibility, data and detailed methods. BPA states that the study, methods and data are not reproducible. We disagree with the BPA statements.

All of the data, detailed methods and mathematical derivations are available. The attached (attachment 1) email documents that on June 12, the FPC received a request from staff of Jones & Stokes, reviewing the Ten Year Retrospective Report under contract with BPA. On June 13 the FPC, in response to this request, transmitted 61 files, providing the specific capture history input files for each of the 2,413,209 fish included in the ten year report. In addition, in the email response, we indicated that FPC staff are available to answer additional questions to assist the consultants' work. With the input files and the formulas, BPA and or their consultants should have been able to generate the components for the formulas, using the widely available MARK or SURPH programs, and then use those components in the formulas in Appendix B of the draft report, or methods explained in the report chapters. In any case the CSS Oversight Committee and the FPC were available to assist reviewers as indicated in the attached emails.

The BPA comment does not explain how BPA and/or their consultants tried to reproduce results. Consequently, it is difficult to respond to the BPA comment regarding reproducibility. Neither BPA nor their consultants attempted to contact the FPC or the Oversight Committee with questions or requests for additional information. BPA and their consultants neither requested a meeting to discuss their attempts to reproduce results nor explained in their comments what specific attempts they made to reproduce results. As always the CSS Oversight Committee and the FPC are available to discuss the report with BPA and their consultants. All of the specific data and the mathematical formulas have been provided to BPA and/or their consultants, and our willingness to respond to additional questions was indicated. Given this lack of information on what more BPA feels they need, we can't determine how to address BPA's request for additional "transparency".

Missing information

BPA states that information is missing and specifically states that formulas for calculating SARs are missing. This is inaccurate. Specifically, Appendix B of the Ten- year Report includes all of the mathematical derivations for the formulas utilized in the Chapter 3 analysis; these include the formulas for calculating SARs. In addition, Chapter 3 includes the formulas for SARs.

Non-standard modeling practices

We disagree with BPA's contention regarding non-standard modeling practices. We have utilized generally accepted, standard statistical procedures for estimation, model-building and associated analyses. Analyses new to this report are based on extensions to methods developed in referenced peer reviewed literature, and methods and assumptions are clearly spelled out. The CSS ten year report is being peer reviewed in this process and CSS products have been peer reviewed in previous years,

Inability to reproduce results

BPA or their consultants' inability to reproduce results do not reflect on the scientific rigor or analytical procedures, modeling or methods used in the Report but perhaps problems with BPA and or consultants attempts. BPA has not described the process used to attempt to reproduce the CSS results, nor did they describe what specifically they were trying to reproduce. They have not availed themselves of the offer by the CSS Oversight Committee to provide guidance or answer questions. All of the input files were available to them and all of the mathematical derivations and formulas were provided in the report.

Latent Mortality

We found a difference in instantaneous mortality rates between SARs of Snake River and downriver wild spring/summer Chinook populations, similar in magnitude to that estimated previously in published literature from spawner-recruit data. The level of differential mortality was relatively small only between upriver and downriver hatchery Chinook (as stated in BPA comments). The BPA or their consultants' proposed adjustment to differential mortality has two major flaws. The BPA adjustment is inconsistent with the definition of differential mortality, and it fails to account for passage survival of transported smolts.

Tagging Results

All of the tagging files and the individual capture history records for each fish were provided to BPA consultants as previously documented. All of the resources of the "CSS organization" were offered to BPA consultants. BPA and consultants did not make any contact with the CSS Oversight Committee or technical staff in their undefined attempts to reproduce results.

Upstream downstream comparison

The BPA comment that data do not support an upriver/downriver comparison is not accurate. Differential mortality is estimable from both PIT-tag and spawner-recruit data. The ISAB (2006) recommended incorporation of additional downriver wild and hatchery populations into the comparison. The CSS Oversight Committee concurs with the ISAB recommendation, and has proposed, but not received BPA funding, to PIT-tag additional downriver populations.

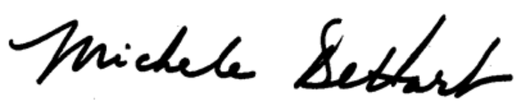
Invalid assumptions

The BPA comment is inaccurate, and the proposed adjustment for in-river migration mortality is inconsistent with the published definition of differential mortality.

Detailed responses to each of the individual comments submitted by BPA are attached (attachment 2).

The CSS Oversight Committee is grateful for the significant investment by BPA in the review and preparation of comments on the draft report. The report has been improved as a result of addressing and incorporating comments. We look forward to future positive collaboration with BPA on future CSS monitoring and evaluation.

Signed



Michele DeHart
Project Leader, Comparative Survival Study

Attachment 1

From: Tom Berggren
Sent: Wednesday, June 13, 2007 2:13 PM
To: Kevin Malone
Cc: Howard Schaller (howard_schaller@fws.gov); Paul Wilson (Paul_H_Wilson@fws.gov); Steve Haeseker (steve_haeseker@fws.gov); Charlie Petrosky (cpetrosky@idfg.idaho.gov); Eric Tinus (eric.tinus@state.or.us); Tim Dalton (Tim.Dalton@state.or.us); Rod Woodin (woodirmw@dfw.wa.gov); Michele DeHart
Subject: RE: CSS Database

Kevin Malone:

Attached is a link to FPC's website from where you may download detection history data used in the CSS. There are 14 directories containing a total of 116 separate data files, which include all wild and hatchery Chinook and steelhead analyzed for SARs in the CSS 10-yr Retrospective Report. This data will be temporarily held on this site until the close of business on June 29, 2007, giving you 12 business days to access and download those data of interest to your review of our draft report. If you have any questions regarding file contents or field names, you may contact me by email.

Tom Berggren

Cc: CSS Oversight Committee members

The CSS file download webpage is at the following link.

http://www.fpc.org/css/css_files.html

From: Kevin Malone [mailto:kmmalone@wavecable.com]
Sent: Tuesday, June 12, 2007 8:08 PM
To: tberggren@fpc.org
Cc: mfilardo@fpc.org
Subject: This is Spam CSS Database

Hi!

I am reviewing the CSS report and would like to get the detection history database for this data set. Specifically, the juvenile and adult detection history for each PIT-Tag used to generate the SAR data etc.

You can send it via e-mail as a zip file or if you point me to a FTP site that would be great!

Thanks!

Kevin Malone
Jones and Stokes

Attachment 2
General Comments

Another aspect of the report used parametric models to partition total variance of metrics into natural variation and measurement error. However, the assumption, for example, that SARs are binomially distributed is inconsistent with the mark-recapture models used to estimate the values. Underestimating sampling error will positively bias estimates of natural variation. The report needs to use goodness-of-fit tests to assess the model assumptions and compare their parameter estimates with those of the nonparametric variance component formulas provided. Their inferences concerning natural variation do not take into account their own findings on ambient effects, the historical distribution of those factors, or how influences such as global climate change many affect projections in the future.

Response: See responses to specific comments on Chapter 4.

Additionally, the CSS incorrectly claims to have addressed the question of whether smolt transportation compensates for effects of the FCRPS on survival of Snake River Chinook and steelhead. At most, the comparison of the SARs of transport fish and inriver fish indicates whether transportation is a viable management option; it is not equivalent to comparing transportation to migration through the unimpounded river. The question of the effect of the FCRPS on salmonid migration and survival is important. However, it is not addressed by the analyses presented in this report.

Response: A major goal of CSS, estimation of the efficacy of transportation, will be described more explicitly in the revised report. In brief, both the absolute realized SARs under the current system, and the ratio of transport SARs to in-river SARs are estimated. Combined with information derived from other sources, it's possible to gain insights on the effect the hydrosystem has had on life-cycle survival rates. It's true that comparison of life-cycle survival under transportation to migration through the unimpounded river cannot be made using information derived only from CSS. However, key components of the comparison include a parameter reflecting any delayed mortality due to transportation (*D*), recent in-river survival rates, and estimates of the proportion of fish transported under recent conditions. These parameters are estimated in CSS and these parameters have been used in models to compare different strategies, including a "dam breach" or "natural river" option (e.g. Peters and Marmorek 2001; Wilson 2003; Zabel et al. *in press*).

Chapter 2

$\hat{Z} = \frac{-\log(\hat{S})}{FTT}$ *(Eqn. 3) provides a convenient but biased estimate of the instantaneous morality rate. Properly, the maximum likelihood estimate of Z would be based on the likelihood model* $L = \prod_{i=1}^{N} Ze^{-Zt_i}$ *and estimator* $\hat{Z} = \frac{1}{\bar{t}}$ *[Eqn. 4] where* t_i *= lifetime for the ith fish* $(i = 1, K, N)$. *However, PIT-tag data do not provide lifetimes for the fish, only travel times for the survivors. Therefore, the PIT-tag data are incapable of estimating instantaneous mortality rates. Any* *relationship between the true estimates of Z [Eq. (4)] and that used in the report [Eq. (3)] may be appropriate at best and seriously biased as worst.*

Response: Our estimates of *Z* (Eqn. 3, above) are the maximum likelihood estimates for *Z* (Seber 1982:216). Contrary to this comment, PIT-tag data not only provide data on the travel times of surviving fish, they also provide survival rate estimates for release cohorts through the CJS methodology that can be used to estimate *Z*. We agree that the estimator suggested (Eqn. 4, above) cannot be used to estimate *Z*, and find this comment to be a useless suggestion in this application. We agree that our use of Eqn. 3 is appropriate, but disagree that this maximum likelihood estimator of *Z* (Seber 1982:216) is seriously biased.

The report seemingly takes a shotgun approach to the analysis. In the results section, which weighing scheme and why its selection was not revealed. The weight selection should be objective.

Response: We used standard statistical methods in the analyses and objective criteria for model building and variable selection. The weighting scheme was objectively determined by the scheme that maximized the adjusted R^2 values for the predictions on the arithmetic scale. The weighting schemes chosen are provided in the tables describing the models evaluated.

Proper weighting should be inversely proportional to the variance except when the variance estimates is correlated with the response variable. In this case, the weight should be inversely proportional to the variance but adjusted to eliminate the correlation.

Response: We evaluated this suggested weighting approach, but found that it resulted in lower adjusted R^2 values than the other weighting methods we investigated.

The report states, "we examined the sign of the parameter coefficients for plausibility and eliminated models with implausible sign." This is a dangerous and potentially misguided approach to modeling. First, such an approach eliminates the possibility that new insights might be developed and assumes all preconceptions are correct. Secondly, it is unwise to directly interpret the sign (+ or −) of partial regression coefficients (Neter et al. 1996:290-291). Such signs do not necessarily indicate a positive or negative relationship between dependent and independent variables but, instead, adjustments of the model in the presence of other covariates. This unorthodox model strategy can lead to odd modeling results (see comments below).

Response: We eliminated the approach of examining parameter signs and now report model fit statistics for all models that were evaluated. As a matter of clarification, this section of Neter et al. (1996:290-291) is primarily focused on the effects of multicollinearity and does not indicate that it is unwise to interpret the sign of multiple regression coefficients.

The report states, "models were fit and ranked according to their AICC and BIC scores." However, many tables (e.g., Tables 2.7-2.11, 2.13) report AIC scores while other tables (e.g., Tables 2.12, 2.15-2.16) report AIC and AICC scores. What was actually done and reported needs to be clarified. For example, are the AIC values in Tables 2.7-2.11 actually AICC and "AIC" is a typo?

Response: All tables with model fit statistics provide the AICc, BIC, R^2, adjusted R^2, delta AICc and Akaike weights (w_i) for each model evaluated.

"Integrated models of fish travel time and instantaneous mortality, with each component modeled being a function of environmental covariates" are mentioned but never described. If a multivariate computational model was actually used, it needs to be provided, along with associated assumptions (providing Eq. 2.2 is inadequate).

Response: We provide equation forms, model fit statistics, and parameter coefficients for the models characterizing median FTT, Z, and S.

Julian day was found in several instances to help describe regression relationships. The implication of this covariate in the models must be described for it is unlike the other covariates considered (e.g., WTT, percent spill, etc.). Julian date is a surrogate for numerous factors that may have a within-season trend including smoltification, flows, temperature, turbidity, etc. If the purpose of the regression analyses is to describe environmental and hydrosystem factors affecting fish response, inclusion of Julian data obscures the results. In some instances, (e.g., Table 2.15-2.16), it does a very good job all by itself!

Response: We provide a description of possible seasonal effects that the Julian day covariate may be capturing. The use of Julian day as a covariate to capture seasonal effects is a common modeling strategy with these data (Berggren and Filardo 1993, Smith et al. 2002, Williams et al. 2005). However, these possible effects (smoltification, photoperiod, fish length/size, predator abundance/activity) are those which are *not* already captured by the other variables examined (flows, temperature, turbidity).

AIC scores cannot be compared across different data sets (Burnham and Anderson 2002:80-81). Comparison of models of FTT and instantaneous mortality versus direct survival is inappropriate and Table 2.2 should be eliminated from the report.

Response: We used the same data set (observed survival rates) to compare with the predicted survival rates (predicted using three different approaches) using AIC values (Burnham and Anderson 2002:63). The table referred to has been expanded.

The authors are totally misinterpreting their estimates of instantaneous mortality Z. In this paragraph, they are equating Z to probability of mortality which is wrong.

Response: For values of $Z \leq 0.1$, mortality rates and Z estimates are approximately equivalent (Ricker 1975). However, to clear up any confusion on the trivial differences between the two, we have provided both daily percent mortality estimates and Z estimates.

The symbolism for box and whisker plots is not universally consistent or known. Captions should explain the symbolism.

Response: Box and whisker plots have a consistent definition and are an elementary topic commonly covered in rudimentary statistical methods courses. The first box and whisker plot now contains a description of what a box and whisker plot represents, for those who are unfamiliar with basic statistical concepts and data descriptions.

Caption fails to indicate which models the results refer to.

Response: The caption now indicates that the survival predictions are based on the variable *Z* approach described.

Omit because AIC are not comparable across different datasets.

Response: The same data set (observed survival rates) was used to judge the different approaches for predicting survival rates using AIC values. In addition, we used root mean squared error, R^2 values and the number of estimated parameters to judge the accuracy of the different survival modeling approaches.

Captions are inadequately described. Symbols for models are cryptic and need to be explained for clarity of interpretation.

Response: Table captions now provide a full description of the symbolism for the variables examined.

The selection of models examined is at times eccentric: Models may include an interaction term without one or both of the main effects included. Purpose of an interaction term is to modify the main effects; it is unclear what the interaction term means in the absence of the main effects.

Response: Models with an interaction term now include both main effects, even though better fits were obtained by omitting one of the main effects in some cases, as was shown previously.

The selection of models examined is at times eccentric: Higher-order polynomial terms are included in models without corresponding lower-order terms, which is not conventional in linear models; for example, squared term without the linear term.

Response: Models with second-order terms now include single-order terms, even though better fits were obtained by omitting single-order terms in some cases, as was shown previously.

Wonder whether this nonconventional approach to modeling is a direct consequence of dropping factors that are perceived to have the wrong sign for the partial regression coefficient (see comment above).

Response: We eliminated the approach of examining parameter signs and now report model fit statistics for all models that were evaluated.

The 20-day curve should be eliminated because the model is extrapolated beyond the range of the data. Fig. 2.1 indicates water transit time in LGR-MCN rarely if ever reaches 20 days.

Response: Water transit times were near or exceeded twenty days for much of the migration season in 2001. As such, the predictions in the figure are bounded by the observed range in the data.

Chapter 3

Page51 (lines 24-26, 33-36) and tables 3.2 (page 63) and 3.4 (page 74) – BPA Comment: Hydrosystem survival and system survival:

Response: In describing both hydrosystem and system survivals, it's clearly indicated that they aren't actual survival rates, and can exceed one. We disagree that they aren't useful in analyzing management options. Hydrosystem survival contains every hypothesized effect on overall survival of any particular proposed hydrosystem action in one term, and can be quite useful in modeling and simply in comparing expected changes in population growth rate due to management efforts in the hydrosystem. We do agree, however, that they aren't really used in the report, and since they can cause confusion and controversy among some readers, we will remove description and estimation of both quantities from the report. Since estimates of pathway probabilities then will not be used at all in Chapter 3, we will move description and estimation of these to Chapter 4, where they are used (for wild smolts).

Page 51 (lines 30-42) – BPA Comment: Assumption of density-dependent mortality needs more support and should be included here.

Response: This assumption related to system survival, which has been deleted from the 10-yr report.

Pages 61-79 (Part A) – BPA Comment: Bootstrap confidence intervals are not superior to theoretical normal theory confidence intervals arising from mark-recapture data analyzed with the CJS model.

Response: If we were only computing estimates of reach survival rates and collection probabilities, there would not have been the need for bootstrap confidence intervals and we would have simply used the theoretical normal theory confidence intervals. However, these parameters which we obtain from the CJS model are only components of more complex parameters. The estimation of number of smolts in categories T_0 and C_1 in LGR-equivalents uses CJS estimates of parameters S_2 and S_3 to expand LGS and LMN detection data, respectively, to starting values at LGR, while category C_0 uses estimates of parameters S_1 in addition to S_2 and S_3 in the estimation of starting smolts numbers at LGR. The estimates of smolt numbers in each study category are effectively combinations of the CJS estimates of S_1, S_2, and S_3 with tallies fish in cells of the reduced m-matrix, which are then divided into the tally of returning adults to obtain the study-specific SARs of SAR(T_0), SAR(C_0), and SAR(C_1). The ratio of SAR(T_0)/SAR(C_0) is used to estimate *TIR*s, and *TIR*s are multiplied by the ratio of S_R/S_T to arrive at *D*. Each of these computed parameters are a more complex function than the starting reach survival components produced with the CJS. The purpose of using bootstrap methods was to produce confidence intervals for these more complex parameters of interest.

Pages 61-79 (Part B) – BPA Comment: Show confidence intervals on all performance measures:

> *i). Geometric means of observed SARs, TIRs, or D values over years of study.*
>
> Answer: In the tables with SAR for each study category, the arithmetic mean and standard deviation is shown (not geometric mean, see the histograms of SARs presented in the 2006 CSS Annual Report), while those of *TIR* and *D* are geometric means. In each table, we will add parametric 90% confidence intervals about the average shown for parameter. Since parameters *TIR*, S_R, and *D* are log normally distributed, we will show confidence interval based on the anti-log of the arithmetic mean and confidence intervals of natural log transformed *TIR*, S_R, and *D*.
>
> *ii). Annual estimates of system survival estimates.*
>
> Response: This parameter is no longer presented in report.
>
> *iii). Annual extrapolated estimates of inriver survival (SR) from LGR to BON (Table D-21 to D-28):*
>
> Response: We will show the estimated 90% confidence intervals for the years with extrapolated estimates of SR with the caveat that those 90% confidence intervals may be narrower than what would have occurred if no extrapolation had been required.

Pages 61-79 (Part C) – BPA Comment: Bootstrap confidence intervals do not easily yield confidence intervals or stand errors on performance measures that are functions of other *parameters. Rather than report measures without some accompanying measure of uncertainty, standard errors or confidence intervals should have been computed in some way.*

Response: With regard to the first sentence of this comment, it appears the reviewer did not understand how the bootstrap process was implemented. Given a release of N fish, each iteration of the bootstrap process was a random draw of N fish with replacement that created a new population of N fish for which all parameters of interest were computed. This process was repeated 1,000 times creating a distribution of 1000 observations for each parameter of interest. This distribution was sorted in order of increasing value, and the parameter value in positions 50 for lower limit and 951 for upper limit were selected for the 90% confidence interval. This approach does readily yield confidence intervals (as well as bootstrap standard errors), so it is unclear why the reviewer thinks bootstrap approaches "do not easily yield confidence interval or standard errors on performance measure that are functions of other parameters."

We are unaware of reasons why the bootstrap cannot be used to estimate confidence intervals (CIs) of quantities that are functions of other estimated quantities. It is true that standard errors of geometric means are easily calculated. However, it's not straightforward to estimate CIs of the geometric mean for short time series. In the special case where number of data points (years, in this case) is 1, the CIs will be lognormally distributed around the geometric mean. With many years of data the CI of the geometric mean approaches a symmetric (*t*-) distribution. However, with the short time series in the present analyses (6-10 years), the confidence intervals of the geomean are neither lognormally nor symmetrically distributed. We have not yet tried to develop an analytical method to estimate CIs of the geomean for short time series. Simulations could be used to estimate CI of the mean, however.

With regard to the second sentence of this comment, we do present the standard errors for the arithmetic means in the tables of annual SARs by study category. It was only in the tables with *TIR*, S_R, and D that we showed only the geometric mean. In our revision, we will show the 90% confidence intervals around the arithmetic mean or geometric mean as is appropriate for the specific table. See our responses to the above Part B portion of this BPA comment for additional details regarding this revision.

Pages 61-79 (Part D) – BPA Comment: Significant differences in point estimates are incorrectly based on non-overlapping 90% bootstrap confidence intervals. The reviewer states that significant differences may still occur even when two estimates have overlapping confidence intervals due to correlation between the two parameters as well as heterogeneity of variances between estimates of the two parameters. The review states that rather than look at the difference between SAR(T_0) and SAR(C_0), we should focus on their ratio TIR as the appropriate measure. The reviewer goes on to state that the determination of significant differences should be recalculated based on formal statistical test, and not on whether confidence intervals overlap.

Response: The review brings up valid points regarding correlation and heterogeneity of variance between the two parameter estimates, and states the *TIR* is the appropriate measure. In the report, we did not confine our investigation of significance to only differences between SAR(T_0) and SAR(C_0), but also indicated that when the lower limit of the *TIR* was greater than 1 there was evidence to statistically demonstrated significance higher SAR(T_0) than SAR(C_0). Based on the reviewers comments, we will revise the text to use the criteria of lower limit of non-parametric 90% confidence interval exceeds 1, which is effectively a statistical one-tailed

(α=0.05) test of H_0 $TIR \leq 1$ versus H_A $TIR > 1$ as the primary measure of whether SAR(T_0) is statistically greater than SAR(C_0).

Page 54 (lines 34-41 and 58-59) – BPA Comment: When inriver reach survival is not directly estimated to BON, you should use the term "extrapolated" instead of "expanded" since you are truly extrapolating past the available data. Did you looked at "per dam" extrapolation in addition to "per mile" extrapolation" You need to add standard errors or confidence intervals to the estimates of extrapolated S_R.

Response: The text will be revised to use the term "extrapolated" instead of "expanded" as recommended. We did compute extrapolations based on "per dam" as well as "per mile," but settled on "per mile" as the more appropriate method. In the reaches between LMN tailrace and MCN tailrace (76.6 miles) and JDA tailrace and BON tailrace (65.86 miles) there are two dams, so the two approaches produce similar results. However, in the MCN tailrace to JDA tailrace (73.94 miles) there is only one dam and a distance similar to the other two reaches noted above. Given this disparity between distances and number of dams involved, we believe the "per mile" extrapolation is more appropriate. We have added confidence intervals to the estimates of extrapolated S_R in Appendix D (as previously stated in responses to BPA Comments on pages 61-79 parts B and C).

Page 58 and 63 (lines 16-18) – BPA Comment: CSS includes steelhead jacks in SAR computation due to steelhead jacks having a fairly stable rate of return, while not including Chinook jacks in SAR computations due to Chinook jacks having a variable return rate. Removing jacks from the analysis because of their questionable contribution to spawning is understandable, but not because of a "highly variable jack return rate.

Response: The CSS report does state that the highly variable Chinook jack return rate among the various hatcheries versus low rate among wild Chinook was one reason for not including jacks in the SAR computations. The other reason, not stated though, is that jack Chinook are considered as having very limited contribution to spawning. We agree with the reviewer that our original sentence about the variable Chinook jack return rate seems out of place, and have deleted it from the text. However, we did not make any statement about steelhead jacks having a fairly stable rate of return. Instead, we simply stated in the methods section that we used 1-, 2-, and 3-ocean returns of steelhead. We will modify the methods section to say "Chinook jacks are excluded due to limited contribution to spawning."

Page 58 (lines 16-18) – BPA Comment: Conclusions (about D averaging 0.5 for hatchery and wild Chinook in recent years) are being presented pre-maturely and inaccurately in the methods section; and that these statements belong in the discussion section with corrections and justification. The reviewer points out that only 3 of 36 point estimates of D were $\leq 50\%$ for hatchery Chinook in tables D-22 through D-26.

Response: We agree with the reviewer that the sentence about D averaging 0.5 does not belong in the methods section. Also, the statement that D was averaging 0.5 applied to wild Chinook only. The 10-yr geometric mean (excluding 2001) was 0.49 for wild Chinook with point estimates in 6 years below 50% and 5 year (including 2001) above 50%. The text will be corrected to reflect this change, and moved to the results section.

Page 58 (lines 26-34) – BPA Comment: Measures S_R and S_T are called "hydrosystem survival," but these are not the hydrosystem survival described on pages 51, 59, and 60. Review wants our intentions explained or clarified.

Response: The "hydrosystem survival" as described on pages 51, 59, and 60 has been deleted from the report. With regard to the measures S_R and S_T, the text will be modified to state: "Therefore, to estimate $SAR_{BON\text{-}to\text{-}LGR}$ from $SAR_{LGR\text{-}to\text{-}LGR}$ for inriver migrating and transported fish, the effect of mortality through the hydrosytem must be removed by factoring out the survival rate from LGR to BON (S_R) for inriver migrants and survival rate in the barge adjusted for the inriver mortality incurred in order to reach transportation sites below LGR for the transported fish (S_T, see Formula 3.10 below)."

Page 59 (lines 13-21) – BPA Comment: Measures SAR_{T1}, SAR_{T2}, SAR_{T3}, SAR_{C1}, SAR_{C2}, and SAR_{C3} need formal definitions, both verbal and mathematical. Also, new notation C1, C2, C3, T1, T2, and T3, is used and needs definitions.

Response: Since we will drop the presentation of hydrosystem and system survival, these quantities will not be used in Chapter 3 and will be deleted. In Chapter 4, where these quantities will still be used, we will clarify their notation and description.

Pages 61-78 and Appendix D (Part A) – BPA Comment: Neither the actual numbers of tagged smolts transported from each dam nor the sample sizes used in the analyses are reported. The review states that this information is necessary for a complete and accurate peer review.

Response: The number of PIT-tags released in each year by species and rear type are presented in Appendix D. Tables D-1, D-3, and D-4 have a column labeled "total PIT-tags" which shows the total tag release each year and analyzed in the CSS for wild Chinook, wild steelhead, and hatchery steelhead, respectively. The actual number of PIT-tagged fish transported are included in Tables D-45 through D-47.

Pages 61-78 and Appendix D (Part B) – BPA Comment: It is unreasonable to assess the effectiveness of transportation based on small transport groups, even if they are augmented by the LGR equivalent approach.

Response: It must be noted that expanding the number of PIT-tagged fish released from LGS and LMN by the in-river survival rates between LGR and those downstream sites is not done for the purpose of augmenting the total transport number. It is necessary when indexing both transported fish and in-river migrants from LGR to expand the downstream counts to account for the fact that some fish die in route to the downstream transportation sites. As shown in Ryding (2006, see Appendix C), there is the need to properly apportion the mortality occurring between LGR and the downstream transportation sites to the transport and inriver study groups in order to obtain unbiased estimates of *TIR*s. We realize that small transport groups limits the ability to show significant differences between transported fish and in-river migrants in many years when the goals of researchers was to return all PIT-tagged fish to the river at the transportation sites. But comparing trends between transported fish and inriver migrants over the years is providing evidence of the level of effectiveness of transportation as a mitigation tool for increasing SARs for wild Chinook and steelhead. The Chapter 4 methods explicitly deal with the effects of small sample size. They produce an estimated mean weighting by sample size, thereby accounting for small sample size.

Pages 61-78 and Appendix D (Part C) – BPA Comment: You should show project-specific TIRs; they are used in estimation of annual SAR in the body of the report, but are not specified.

Response: We assume that the second part of this comment applies to the annual estimates of overall SAR reported in this report. The overall SAR is computed by taking the study-specific SAR of groups T_0, C_0, and C_1 and weighting these SARs by the estimated proportion of fish in the total population (untagged and tagged) represented by each study-specific SAR. There is no use of *TIR* in this estimation. The results of evaluations of project-specific *TIR*s are covered in Chapter 4.

Page 61 (lines 26-28) and page 68 (lines 14-15) – BPA Comment: It is unreasonable to say that 2004 SAR is "low" at this point, since the 2004 returns are incomplete.

Response: With 3-ocean returns accounting, on average, for about 30% of the total adult return, and the SAR for the 2004 wild Chinook (based on the 2-ocean return) estimated at 0.30%, 0.31%, and 0.18% for categories T_0, C_0, and C_1, it was obvious that even when the 3-salt returns are added, the resulting complete return will provide "low" SARs.

Page 61 (lines 30-32) – BPA Comment: A reference made in the ISAB review of the 2005 CSS Annual Report refers to the NOAA finding that PIT-tagged survival is less than untagged survival. If the NOAA finding is true, then comparing SARs from PIT-tagged fish to target values is unreasonable unless we know the size of the bias introduced by tagging or tag loss.

Response: We address this issue in detail in Chapter 5. This line of reasoning assumes that the run reconstruction approach is correct. However, it may be that the difficulties in applying that approach has created SAR estimates that are too high. The "true" population SAR may lie somewhere between the levels estimated by these two methodologies.

Page 62 (line 13) and Table D-21 – BPA Comments regard the use of geometric mean to summarize point estimates of SAR, TIR, and D across years.

> *i). Use of the geometric mean needs justification, especially considering past criticism and the fact that the geometric mean will always be lower than the arithmetic mean.*
>
> Response: This same comment was made by BPA on the 2006 CSS Annual Report. The response was given on pages 170-171 of that document. In general, SARs for each study category approximate normality, as do the individual reach survival rates computed by the CJS method. However, the parameters S_R (i.e., the product of $S_2 \cdot S_3 \cdot S_4 \cdot S_5 \cdot S_6$), *TIR*, and *D* each appear to be lognormal distributed with skewness to the right. For these reasons,

the arithmetic mean was used for parameter SAR and the geometric mean was used for the other log-normally distributed parameters.

The geometric mean is a better measure than arithmetic mean of central tendency for right skewed (log-normally distributed) distributions such as *TIR* and *D*. They both represent ratios of survival rates, for which the ordering (i.e. which is numerator and which denominator) is arbitrary. From Zar (1984, p. 24): "[The geometric mean] finds use in averaging ratios where it is desired to give each ratio equal weight".

ii). Standard errors or confidence intervals need to be reported for the geometric mean (see earlier comment and suggestion.

Response: We have added standard errors and confidence intervals to the geometric means presented in Appendix D. However, the methods used for calculating these confidence intervals with short time series may not be appropriate, as discussed in response above.

iii). Low precision on D and TIR casts doubt on conclusions based on the geometric mean, especially those based only on a point estimate.

Response: We agree that low precision on annual estimates of *D* and *TIR* suggests that an unweighted mean should be interpreted cautiously. However, in the presence of large differences in mean values from target values, some inferences may be in order. The variable precision among annual estimates was a prime motivating factor in applying the methods used in Chapter 4, which allow stronger conclusions about the central tendencies of these quantities. As the number of years increase, the precision of geometric means will improve.

The reference to some estimates being only point estimates appears to refer to the parameter S_R and not *D* and *TIR*. As stated in earlier responses, we will show the 90% confidence interval for those S_R values that were extrapolated from a shorter reach, with the caveat that these confidence intervals will generally be narrower than would have occurred if sufficient data had been available to directly estimate the reach survivals in those lower reaches affected.

iv). The geometric mean inherently dampens the effect of extreme values, so the policy of excluding 2001 values from the geometric mean needs further justification.

Response: Excluding 2001 from the geometric mean was not a policy action. The drought conditions of 2001 were so unlike the other years that it was of interest compare the resulting *TIR* and *D* estimates of 2001 to the geometric mean of the other years. Data from 2001 were included in all estimates of *TIR* and *D* distributions made in Chapter 4

Page 66, 70, and D-17 (Tables D-29 and D-30) – BPA Comments: Annual SAR.

i). Annual SAR is discussed often and is described in word, but is never defined formally. An equation is needed to see exactly how the various components are incorporated.

Annual SAR values should be reported in a table for all species and stocks, with confidence intervals or standard errors.

Response: Coverage of the SARs described on pages 66 and 70 has been moved to Chapter 4, where equations and tables of results are presented.

ii). It would be useful to compare the annual SAR values to a simple ratio of the number of adults at LGR divided by the number of juveniles at LGR

Response: We disagree because the study fish do not migrate through the hydrosystem via the different routes in the same proportion as the untagged fish. Therefore weighted SARs are necessary.

iii). Tables D-29 and D-30 – BPA Comments that these tables should be explained clearly in text, using precise equations and clear definitions. It is unclear how the values reported here were defined, estimated, and used to compute the annual SARs. It is unclear what the S's mean, and what reaches they apply to. It is unclear where the covariances come from. No comparable tables were provided for hatchery fish.

Response: Appendix D presents information relevant to the whole document, not just to Chapter 3. These tables refer to work presented in Chapter 4, not Chapter 3. Apparently the commenter assumed they described an analysis in Chapter 3. Table D-29 is referenced in Chapter 4, and nowhere else. A reference to Table D-30 was inadvertently omitted from Chapter 4 and has been added. The purpose of the tables is clearly labeled in their captions; namely, to estimate covariance between pathways to estimate overall SAR mean and variance. The exercise was performed only for wild fish. The *S*s are also clearly defined in the captions. Moving the pathway probability language from Chapter 3 to Chapter 4, where the tables are referenced, should make the purpose of the tables obvious.

Page 67, Figure 3.7; page 70, last paragraph – BPA Comment: Figure 3.7 shows that the trend in SAR for wild fish over 2- or 3-yr time periods mimics the trend in SAR for certain hatchery stocks. However, Figure 3.7 also shows that SAR for wild fish did not closely track SARs for any single hatchery throughout the entire time period considered. It is therefore uncertain which single hatchery could be used as a surrogate for wild fish in future years. Also no error bars are provided on Figure 3.7.

Response: We agree that no one hatchery mimics the trend in overall SAR of wild Chinook, nor for trends in S_R (Figure 3.8), *lnTIR* (Figure 3.9), and *ln*D (Figure 3.10). That is why we do not make any recommendations for using only one hatchery as a surrogate. As for the lack of error bars in Figure 3.7, we present the 90% confidence intervals in Appendix E for the overall SAR parameter as "tot_sar." This appendix was not available at the time of the review. With up to 5 to 6 curves shown in Figure 3.7 across the years, the inclusion of error bound on each would have been too cluttered.

Pages 67 to 78 (Figures 3.7, 3.8, 3.9, 3.10, 3.13, 3.14, 3.15, 3.18, 3.19, and 3.20) – BPA Comment: Confidence intervals or standard errors are needed on these graphs.

Response: The goal of the figures was only to show the trends across years for the groups of fish being compared, and not to test whether significant differences occurred. We present 90% confidence interval in Appendix E for the overall *SARs*, S_R, untransformed *TIR*, and untransformed D. The 90% confidence intervals may also be found in Appendix D tables D-21 through D-28 for S_R, *TIR*, and *D*. Since the goal of the plotted data in these figures was aimed at only comparing trends over years, the error bounds about each curve was omitted in order to keep the plot uncluttered.

Chapter 4

Pg 81, 3rd paragraph – As the SARs are calculated in Chapter 3, they certainly do not have a binomial sampling variance, for both the numerator and denominator (i.e., C0 fish) are estimated random variables. For a binomial variance to be true, the denominator of the SAR would have [to] be known without error

Response: The numerator for SARs of any group is number of adult returnees detected at LGR, which is a count and not a random variable. It's true that the denominator for SAR of C_0 is an estimated quantity; however, as indicated later in Chapter 4, the CVs are small, and as demonstrated below, the deviation in variance from a true binomial is minimal. Similar methods of removing binomial variance from survival rate estimates which are not strictly binomial processes have been used. For example, Morris and Doak (2002) present an example using Kendall's (1998) beta-binomial method with data from desert tortoises: "[T]he capture-recapture method used to estimate survival doesn't yield a directly observed sample size. Instead, we used a rounded estimate of the total number of individuals that would have produced the observed number of live tortoises seen at the end of each time period, given the estimated survival rate" (pgs. 266 and 270.)

The variance of the ratio of returning adults to estimated number of smolts can be derived using the delta method, assuming both the numerator and denominator are random variables. A close approximation of the variance of the ratio of two random variables X and Y is (after Blumenfeld 2001, Eq 2.29)

$$Var\left(\frac{X}{Y}\right) \cong \left(\frac{\mu_X^2}{\mu_Y^4}\right)\sigma_Y^2 + \frac{\sigma_X^2}{\mu_Y^2} - 2\left(\frac{\mu_X}{\mu_Y^3}\right)\rho\sigma_X\sigma_Y,$$

where μ and σ^2 are mean and variance, respectively, and ρ is the correlation between X and Y. In the true binomial, variance of Y is zero, and the variance of the ratio reduces to the usual formula for variance of a binomial proportion p, i.e. $p(1-p) / N$, where N is the number of trials (number of smolts). By plugging in a value for CV of N when N is not known with certainty, the expected true sampling variance can be estimated. As noted in the discussion of Chapter 4, CVs of the estimate of C_0 are generally 2-4%. Below, we explore the effect of a CV of 4% in the numerator, along with two assumptions about the correlation between smolt numbers and adult returns (ρ), and two assumptions about mean smolt numbers, which reflect most of the range in annual C_0 estimates. Mean SAR is assumed to be 1%, which is close to estimated values of SAR(C_0) for both wild steelhead and wild Chinook.

Table 1 shows that the effect of observed levels of variance in the denominator of SAR(C_0) is minimal. Simulations of binomial draws from a normal random variable representing C_0 indicate that, as expected, correlation between adult returns and smolts numbers increases with smolt numbers. Even at 5000 smolts, however, the estimated correlation at CV of C_0 = 4% is only 0.27, suggesting that the actual sampling variance departs little from the assumed binomial variance. Additionally, a positive correlation between smolt number and adult returns results in the binomial variance *over*estimating the sampling variance. This suggests that assuming binomial sampling variance may result in slight *under*estimation of environmental variance, for the range of correlations pertaining in this analysis. An expanded version of this analysis has been added to Chapter 4.

Table 1. Effect of CV of 4% in C_0 estimate on sampling variance of SAR(C_0), for different correlations and mean smolt number. SAR assumed = 1%. Binomial variance was assumed in Chapter 4 analyses. CV of SAR is sqrt (variance) / 1%.

Mean C_0	ρ	Actual variance	Actual CV	Binomial variance	Binomial CV
200	0	4.97×10^{-5}	70%	4.95×10^{-5}	70%
200	0.5	4.68×10^{-5}	68%	4.95×10^{-5}	70%
5000	0	2.14×10^{-6}	15%	1.98×10^{-6}	14%
5000	0.5	1.58×10^{-6}	13%	1.98×10^{-6}	14%

Page 82, lines 15-17. Akcakaya (2002) is cited as a foundation for the method used to remove sampling variance to estimate environmental variance. The method presented in Akcakaya (2002) is appropriate for census data, but not for mark-recapture data, such as the data analyzed in this report. Akcakaya (2002) refers to both Burnham et al. (1987) and Gould and Nichols (1998) for variance-components method of removing sampling variance from mark-recapture data (see below, comment on pages 82-87).

Response: Gould and Nichols (1998) point out that if the population parameter is known (directly observed), there is no variance component associated with sampling error. Gould and Nichols' analysis considered cases with two sources of "sampling" variability (pg. 2532): 1) variation associated with the inability to count at sampling period $i + 1$ every marked survivor from period i, and 2) demographic stochasticity producing binomial variation in the number of marked survivors at the end of period $i + 1$. In the present analysis, there is no sampling variance of the first kind. All (or nearly all) surviving adults are "captured" by PIT-tag detection at LGR, i.e. there is a "census" of survivors. Therefore, since the present analysis deals only with demographic stochasticity, the more involved methods of Gould and Nichols for estimating the first kind of sampling variance and its covariance with the second kind are not required.

Page 81, 4th paragraph – The belief that there is a single probability distribution of SAR, TIR, or D over a long time period assumes that there is no temporal trend in the measures, such as may *be caused by global climate change. Chapter 3's focus on trends in these measures suggests an assumption that the measures are changing over time, which is inconsistent with the assumption that they arise from a single beta distribution*

Response: CSS's primary purpose is in data collection and monitoring, and in particular estimating SARs and the efficacy of smolt transportation. Using the presents methods to estimate distributions reflecting inter-annual variation in SARs and their ratios observed in the recent past requires no beliefs about the factors influencing SARs. In the introduction of Chapter 4, the assumption under which the distributions derived would be useful for prospective modeling is clearly stated (pg. 81). In any system, the future cannot be guaranteed to be identical to the past, yet there is no end of literature presenting estimates of recent population abundance, survival rates, population growth rates, etc, in an attempt to understand the current state of the system. Describing *what* has occurred is not inherently inconsistent with exploring hypotheses about *why* it occurred.

Page 86, lines 19-22 - Equation(4.4) for the variance of a product applies only for independent random variables. This equation cannot be used to calculate the variance of a product of inriver survivals over adjacent reaches (i.e. S_R,), because these survival estimates are correlated as based on the CJS model. Instead, the delta method (Seber 1982:7-9) should be used.

Response: The paragraph immediately under Equations 4.4 and 4.5 indicates that the assumption of independence of the random variables is required. Here and elsewhere this assumption is made, evidence supporting its reasonableness is provided. Description of the accuracy of the bootstrap procedure in reproducing overall variance in S_R from individual reach survival rate estimates which covary is presented elsewhere in these responses.

Page 82-87 - Kendall's (1998) method is a parametric approach to variance component estimation that makes unnecessarily restrictive assumptions, i.e., a. Measurement error is binomially distributed. b. SARs are beta-distributed (and following equations and numbered points).

Response: The commenter has misunderstood the method of variance partitioning used. As explicitly stated in the Chapter twice (pg. 82-83), Akçakaya's (2002) method of variance partitioning, rather than Kendall's, is used. On page 82 we explicitly note that Akçakaya's method is an alternative to the approach of Kendall. Akçakaya's method involves no assumption about the form of distribution of the resulting survival probability. Our rationale for representing the resulting environmental variances with beta distributions is provided elsewhere in Chapter 4. The goodness of the assumption of binomial sampling error of C_0 SARs is discussed above.

There are several implications of the parametric approach taken to variance component estimation used in the CSS report, including the following: (following bullet points).

Response: See Chapter 4 for rationale for choosing beta distributions to represent variability in SARs. Kendall (1998) and Morris and Doak (2002) use similar methods to estimate beta distributions to describe variability in survival rates. See these authors for more detailed rationale, and survey the literature on stochastic population modeling for numerous examples of using the beta to represent variation in survival probability. Those authors do not expect proof that the limited data in hand in most conservation problems conforms to a beta distribution (this is impossible with the short time period data sets available in most conservation problems—it would take many, many years of data to allow discrimination between beta and alternative distributions). Perhaps the commenter could suggest a different probability distribution that would better reflect variation in survival rates over many years.

The exact form of the beta distribution used is presented in Chapter 4—it is the identical form as used by Kendall (1998), as referenced. Equations 4.2 and 4.3 (4.3 and 4.4 in the revised Chapter) show how the parameters of the beta distribution are derived from the mean and environmental variance derived from using Akçakaya's approach. The commenter's equations 5 and 6 can be derived from equations 4.2 and 4.3 by solving for mean and variance; or the converse operation can be performed.

Where data are sufficient for plausible estimates of correlation, our analyses do not assume that SARs of different groups vary independently. In fact, in estimating *TIR* and *D* distributions for Chinook (where data are sufficient for estimation of correlation), we include covariance between transport and in-river groups. With regard to global warming, see earlier response to page 81 comment.

Page 88, Table 4.1 • The estimated demographic variance is greater than total variance, suggesting something is wrong and thus casting doubt on all methods and results in this chapter. • Observed correlations between point estimates of SAR for transport and C0 groups for wild steelhead are explained by small transport groups and so are not used. However, such small transport groups (we are not told the actual sizes) produce unreliable parameter estimates that can seriously distort interpretation of results.

<u>Response</u>: Gould and Nichols (1998), which the commenter commends, produced negative estimates of variance (due to estimated sampling variance being > total variance) for a number of their sample data sets. They reference literature indicating that negative estimates of variance are not uncommon in the variance components literature (pg. 2534-2535). In the CSS study, the one case of estimated sampling variance slightly exceeding total observed measurement (steelhead transported from LMN) is a consequence of large sampling variation due to only 8 PIT-tagged adults returning to LGR over the 6 years. In this case, a reasonable and conservative approach is to use the observed inter-annual variance as an estimate of environmentally driven variance.

Uncertainty in parameter estimates is explicitly estimated and accounted for in these procedures. The effects of small sample sizes combined with low SARs can be seen in the resultant wide confidence intervals for SARs of LGS- and LMN-transported steelhead (Figure 4.2). The effect of this uncertainty is carried into estimates of *TIR* and *D* for these projects and explicitly presented. Assuming independence of SARs in estimating *TIR* and *D* distributions is a reasonable and conservative default procedure in this case. The relevant raw data, including numbers of PIT-tagged fish transported from each project, and detected upon return as adults at

LGR, can be found in Appendix E of the 2006 CSS Annual Report. The raw data were also provided in electronic form to BPA at the start of the comment period.

Chapter 5

BPA General Comments, p. 2: *The CSS continues in Chapter 5 its comparison of upstream and downstream Chinook salmon stocks. As in the past, multiple upstream hatcheries and collection points are used, while only a single downstream hatchery and collection point (for wild fish) is used, despite the ISAB's recommendation to incorporate more downstream stocks. Given that this is a retrospective report, it is understandable that the CSS could not immediately include additional downstream stocks. While the CSS does perform useful comparisons of biological characteristics of the upstream and downstream stocks, their upstream-downstream analysis is invalid in other critical ways. The CSS uses an invalid performance measure to identify delayed mortality caused by the hydrosystem. This approach assumes no natural mortality for smolt should occur between upstream and downstream sites. When the performance measure is corrected for the longer migration of the upstream stocks, there is little or no evidence of delayed hydrosystem mortality for hatchery Chinook salmon. Similarly, the CSS Report does not consider the longer distance to travel for upstream stocks when comparing travel and arrival times of upstream and downstream stocks. Even if the hydrosystem were not in place, the upstream stocks would still have farther to travel than downstream stocks.*

Response: To clarify, three downriver populations are included as an aggregate in the analysis for wild Chinook: North Fork, Middle Fork and upper mainstem John Day Rivers. We have noted both the ISAB recommendations and the CSS proposals to increase the number of downstream wild and hatchery populations, which BPA has yet to fund.

The BPA reviewers appear to be confused on the purpose of the upriver/downriver analysis, which was stated in the Chapter 5 introduction (p. 106): "*our specific interest … is whether upriver/downriver differences in SARs for wild and hatchery stream-type Chinook were consistent with the differential mortality estimated from SR [spawner-recruit] models for wild populations"* Previous published SR analyses indicated there was a systematic increase in mortality for Snake River populations, which did not occur in the downriver populations, associated with the construction and operation of the FCRPS (e.g., Deriso et al. 2001; Schaller et al. 1999). In the SR model formulations, any differences in smolt mortality caused by different travel distances would be incorporated into the intrinsic productivity (Ricker "a"). Obviously, the migration distance for upriver and downriver populations did not change over the time period of FCRPS development; Water Travel Time (WTT), Fish Travel Time and hydro impacts did change with this development. WTT for Snake River stocks before FCRPS development were only about 2-3 days; Snake River smolts were historically able to arrive at the estuary more in synchrony with their morphological, physiological and behavioral development (e.g., Budy et al. 2002; ISG 1999). Available evidence from a mostly free-flowing migration corridor (Whitebird trap on the Salmon River to Ice Harbor Dam) also suggests smolt survival was high before FCRPS development (Raymond 1979). Applying the survival per mile from the Raymond study the information suggests that the historic survival from Lewiston to Bonneville dam was over 90%.

The BPA reviewers appear to be confusing differential mortality and delayed mortality. The analysis in the CSS report estimated differential mortality based on SARs to compare with differential mortality estimated by SR analyses (see equations 5.2 and 5.3). We did not explicitly estimate *delayed* mortality for in-river migrants, although the upriver/downriver SAR differential mortality comparisons are relevant to such an analysis.

• *On a yearly basis, p should be estimated as (i) Manly-Parr formula, (ii) CJS formula, but not the CSS formula p = (N detected at BOA) / (N detected at BOA + N passing BOA undetected that were later detected upriver).*

Response: From the reviewer's comment, it was apparent that the formula shown in footnote 5 of Table 5.9 caused a misunderstanding of our approach. That footnote has been corrected. Our approach is identical to what the CJS model produces in a three site model – site 1 for release, site 2 for BOA and site 3 for pool of upriver dams. In the Burnham et al (1987) monograph, the estimate of collection efficiency at site 2 is $p_2 = m_2/ m_2 + z_2 \cdot (R_2/r_2)$. With only 2 recovery sites, this equation simplifies to the following using the reduced M-matrix in the Burnham monograph for k=3 sites:

Cohort	Site 1 Release	Site 2 BOA	Site 3 Upstream	Sum detections
1	R_1	m_{12}	m_{13}	r_1
2	R_2		m_{23}	r_2 .
	Column sum	m2	m3	

Sum for z_2 is m_{13}
Sum for r_2 is m_{23}
$R_2 = m_{12}$ since there are no removals at BOA
$m_2 = m_{12}$

Formula for $p_2 = m_2/ m_2 + z_2 \cdot (R_2/r_2)$
$= m_{12} / (m_{12} + m_{13} \cdot (m_{12} / m_{23}))$
$= m_{23} / (m_{23} + m_{13})$

The number of fish in m_{23} = *N* jointly detected BOA & upriver
and number of fish in m_{13} = *M* passing BOA undetected & detected upriver

Substituting these equalities gives the formula that we are now showing in footnote 5 of Table 5.9. Therefore, we are actually utilizing the CJS model approach and producing a valid estimate of p_2 at BOA.

BPA comment: Page 106, lines 11-22
• *Critiques of the single release-recapture (SR) analysis and PATH have demonstrated the reliance of latent mortality results on untestable assumptions, e.g., stock-specific Ricker a's*

versus a common Ricker a. Additionally, climate effects have been shown to account for the majority of latent mortality. These criticisms should be addressed in this chapter.

Response: The BPA reviewers seem to be confusing delayed or latent mortality with differential mortality; also SR is the abbreviation for spawner-recruit, not single release-recapture (see p. 115). The differential mortality estimated from PIT-tag SARs (equation 5.3) can be used ultimately to test differential mortality estimates using different SR (spawner recruit) model formulations. It is important to note that the reviewers are criticizing the published material we referenced, however, we did not perform SR analysis in the CSS report. The purpose of the CSS PIT-tag analysis was to provide independent estimates of differential mortality, for comparison with estimates from published SR analyses (Schaller et al.1999, Deriso et al. 2001, Schaller and Petrosky 2007). We are aware of one alternative SR model that suggests differential mortality may be low, which uses a common Ricker "a" for all populations (R. Hinrichsen, unpublished manuscript); other models investigated by Hinrichsen yielded differential mortality estimates similar to that in Figure 5.16 in the CSS report. Given the 4-fold difference in SARs estimated between Snake River and downriver populations, the common Ricker "a" hypothesis does not appear very plausible. Other issues with this hypothesis include the habitat quality differences among Columbia Basin streams (and thus expected differences in intrinsic productivity) and the fact that the common Ricker "a" formulation produces other questionable parameter estimates. Regardless, by continuing and expanding CSS PIT-tagging of upriver and downriver populations more formal testing will be possible through analyzing these SAR estimates.

BPA comment: Page 106, lines 19-20
• It is not explained and it is unclear how direct mortality, differential delayed mortality of transported smolts, and the common year effect were accounted for in the SR comparisons.

Response: We provided three references which provide detail regarding how delayed mortality of in-river migrants may be partitioned from total mortality. Since we did not explicitly estimate delayed mortality in this report, we did not provide equations from these literature sources that did make delayed mortality estimates.

BPA comment: Page 107, line 26
• "Overall SAR" is never defined, either here or elsewhere in the report. Presumably it is equal to "annual SAR," which is also never defined analytically.

Response: We added the definition of overall SARs and a reference to the detailed analytical description in Appendix B (see page B-10 – *Estimation of overall annual SARs*).

BPA comment: Pages 110-111
• Run Reconstruction SARs: Include jacks and adults; measure returns to mouth of Columbia River.
• CSS SARs: Include only adults (Chinook), no jacks; Measure returns to LGR
• Are run-reconstruction SARs and CSS SARs really comparable? It has not been justified that direct comparison of the measures is appropriate.

Response: We modified the language to indicate that both run reconstruction and CSS SARs in this analysis represented returns to the uppermost dam (Lower Granite since 1975) adjusted to account for harvest. Our initial comparison had the (quantitatively minor) inconsistency that we included jacks in the run-reconstruction estimates, which we have fixed.

BPA comment: Page 112, lines 15-19: How is WTT defined?

Response: We added the following language: Water velocity in the mainstem migration corridor is generally expressed as the average time (in days) it takes a water particle to travel through a river reach (water travel time) during a specified period.

BPA comment: Page 114, lines 3, 9; Figures 5.5, 5.6
• What does "frequently incorporated in multiple regression models" mean?

Response: We changed "incorporated" to "selected" in the caption.

BPA comment: Page 115, Multiple Factor Model, lines 5-31
• How were candidates for independent environmental covariates selected? What were they? Only WTT, PDO, and an upwelling index are named, and it is unclear whether other covariates were considered.
• Harvest and temperature are known to affect SARs and do not appear to have been considered.
• Were any other "inriver" predictors than WTT considered?
• Were interaction terms considered in the multiple regression models?
• Typo in SAS version (presumably 9.1, not 91).

Response: Candidates for independent environmental covariates were those that have been previously linked to, or hypothesized to influence, salmon SARs (p. 112). Other potential juvenile migration variables covariates for future analyses may include a measure of spill proportion and proportion of the run transported. Because SARs in this analysis represented pre-harvest adult recruits, harvest was already accounted for. We did do some exploratory analysis with average monthly sea surface temperatures at various latitudes. However, it was not very informative and we believe sea surface temperature was incorporated by the PDO, a large-scale index of sea surface temperature anomalies in the North Pacific Ocean. We did not include any interaction terms, although, this may be attempted (for the longer time series) in future analyses. The SAS version typo was corrected.

BPA comment: Methods: Snake River and Downriver SAR Comparison (pp. 115-119)
• There has been much previous criticism of the upriver-downriver comparisons made by the CSS and of the spawner-recruit model used to justify the upriver-downriver comparisons. Insufficient response has been made to these criticisms.

Response: We went into detail addressing each off the past criticisms for the upriver-downriver approach on page 119-120. We focused on the published upriver-downriver criticisms and the published responses to these criticisms. In addition, we provide a summary of analyses comparing biological characteristics of the two population groups.

Page 116
• Lines 7-8
− How is μ_t defined and estimated? Provide an equation showing how value is calculated. Is this the same μ as in Eq. 5.3, or is it the differential mortality defined verbally based on Eq. 5.2?
− The "delta model" should be defined.

Response: We did not estimate μ_t from SR data in the CSS report, we only compare PIT-tag estimates of differential mortality to previously published estimates of μ_t. We specifically referred the reader to Deriso et al. (2001) equations 4-6 for estimation of μ_t. The delta model was defined as the primary model in Deriso et al. (2001) just above equation 5.2 (p. 115 line 44).

BPA comment: • Equation 5.3: If there is no delayed mortality from hydrosystem, then we expect $exp(-\mu_{SAR,t}) = S_{J(LGR\text{-}JD)}$. This important point is omitted from the report.

Response: The subscript "J" in the reviewer's comments is not clear to us. However, see our response to the reviewer's table 1. If we understand the reviewer's point, partitioning in-river survival ($S_{(LGR\text{-}JD)}$) from the SARs is not analogous to estimating differential mortality from SR data. Also, this formulation, as we interpret the reviewer's point, does not account for the large proportion of fish which are transported.

BPA comment: Page 117
• Line 18: Only a single hatchery (Carson) is used for the downstream hatchery Chinook salmon.

Response: The CSS study has only received funds to maintain a long time series of PIT-tag SARs for only one downriver hatchery.

BPA comment: Page 119, Table 5.9
• This table is very difficult to understand. The caption does not agree with the notation used in the table. Values reported in the table are not sufficiently explained. It appears that the formula used to estimate BOA detection efficiency (p) is wrong.

Response: Agree that caption does not clearly state the purpose of the table, so it has been revised to read:
" Table 5.9. Estimated PIT-tag detection efficiency of combined adult detectors at Bonneville Dam based on combined unique detections of PIT-tagged adults at McNary, Ice Harbor, and Lower Granite dams."

Also, footnote 5 was misleading as currently stated and has been revised to read:
" Calculated as p = (N jointly detected BOA & upriver) / (N jointly detected BOA & upriver + M passing BOA undetected & detected upriver)

• What are the values reported in the row "GRA, MCA, IHA?"

Response: The sum of unique PIT-tagged adults (≥ 2-ocean) detected at either IHA (Ice Harbor Dam, where IHA and ICH are possible detection site names), MCA (McNary Dam, where MC1 or MC2 are possible detection site names), and GRA (Lower Granite Dam). Each returning PIT-tagged adult is counted only once from this pool of three recovery sites.

• MCN and IHA are not mentioned in table caption.

Response: Caption has been rewritten so that all three dams are included

• The estimate of p based on detections at BON and upstream is INVALID if it is based on detections from different years, unless upriver adult survival to GRA is constant across return years, and detection probabilities at MCN, IH, and GRA are constant across return years. This is not true, so estimates of BOA detection efficiency presented here are invalid.

Response: The annual detection efficiency probabilities at BOA were estimated at the level of the smolt migration year, so as to allow a single expansion factor at BOA for total adult return counts. The reviewer's concern that upriver adult survival and detection probabilities may change across years is not a problem since we are creating the BOA detection efficiency as a conditional probability, given the sum of unique (counted only once) PIT-tagged adults detected above BOA. Since these fish are detected above BOA, we know they were alive when passing BOA, and so a conditional probability calculated as p = (N jointly detected BOA & upriver) / (N jointly detected BOA & upriver + M passing BOA undetected & detected upriver) is a valid approach.

BPA comments: Methods: Comparison of biological characteristics of Snake River and downriver smolts (pp. 119-121)
Page 120
• In general for upstream/downstream comparisons, was goodness-of-fit considered or examination of residuals performed? Show results.
• With only 6 years of data, this is not a long time series, which limits the amount of useful information that can be gleaned from it.

Response: It is not clear what reviewers are suggesting with the first comment. Six years of data are what we have available, however, sample sizes (numbers of tagged smolts) are large enough within and across years to detect statistical differences where they exist.

BPA comments Page 121
• Lines 13-14: No migration distance is given for JDAR1 fish. Comparison of survival and travel time between upstream and downstream fish should incorporate migration distance for the two groups of fish.

Response: Reviewers' comment is not clear; we presented the migration distances in lines 13-14.

BPA comments • Lines 40-41: Basing analyses on (Number of BON detects/Number released at trap) assumes that all groups have the same conditional detection probability at BON. This is likely to change with arrival timing.

<u>Response</u>: It is unclear what the reviewers point is?

<u>BPA comments:</u> Results: Overall SARs (pp. 122-127)
Page 122, lines 32-34
• "Removing sampling variability" resulted in lower mean SAR. Does this always occur?
Page 126, lines 17-19
• The CSS has been using a geometric mean previously, but here does not identify the type of mean used for mean SAR.
• It is unclear what the reference to the t-distribution means. If a formal t-test is being performed, this should be stated simply. Note that while these arithmetic means may be compared using a t-distribution, the geometric mean should not.

<u>Response</u>: The variation portioning ("process error") method used in Chapter 4 uses a weighted mean SAR, which usually will differ from the unweighted mean. The amount and direction by which they differ depends on how sampling error is distributed among years with varying point estimates of SARs.

In the draft report we did not log transform the SARs. In the final draft we recalculated the mean SAR based on natural log transformation and the percent of the distribution above 2%, and modified the text accordingly.

We did formally use a t-test and specifically stated our methods on page 107 lines 33-39.

<u>BPA comments</u>: Results: Relationship between SAR and environmental covariates (pp. 128-131)
Page 128
• Lines 4-8: The data for the PIT-tag SARs and environmental factors are not presented in this report.
• In general, references to figures should be proofread. There are mistakes in figure references throughout the chapter, making it difficult to follow the narrative.

<u>Response</u>: The CSS PIT-tag SARs (LGR-LGR) are in Appendix E (data was sent to BPA reviewers on request). We also cited the source of the run-reconstruction SAR data set, and provided the websites for environmental data. We corrected the figure references in the final draft.

<u>BPA comments</u>: Line 11: What is meant by "bi-variate results?" Is this regression of a single response variable on a single predictor variable? A vector response variable on one or more predictor variables? A single response variable on two or more predictor variables?
• Table 5.4: Did the CSS consider correlation between PDO and UP45n? Both types of measures are used in the same regressions, apparently.
Page 129, Table 5.5
• It should be explained why SepPDO is used rather than JulyPDO as a covariate, when JulyPDO looks better than SepPDO for both the long and current time series. Page 129, Table 5.5

• It should be explained why SepPDO is used rather than JulyPDO as a covariate, when JulyPDO looks better than SepPDO for both the long and current time series.
Page 130
Page 130
• In general for regression with environmental variables:
− What was the set of candidate predictor variables? Was it only PDOs, UP45ns, and WTT?
− How model selection was performed needs to be specified?
• Lines 13-14 – The report says that WTT was "less significant for the shorter time series," but Table 5.7 indicates that WTT was not at all significant if the model includes upwelling index (Table 5.7, Current Time Series).
Page 131, lines 1-2, and Table 5.7, Current time series
• What model selection criterion were used to identify the "best" model?
• The "best" model shows no predictor variable significant at the 10% level when upwelling index is included.
• Without upwelling index (NovUP45n), SNWTT and MayPDO become significant. Was multicollinearity between these parameters and NovUP45n considered? And how?

Response: Our use of the term bi-variate results refers to regression of a single response variable on a single predictor variable. Our primary concern with correlated independent variables was to screen against highly correlated monthly variables within the PDO (such as between April and May or May and June) or the within the upwelling indices. However, the correlation between SepPDO and NovUP45n model selection (ocean variables selected for the best 3 parameter model for the long time series), was negligible (-0.02). JulyPDO was screened out from the regressions because it was highly correlated with MayPDO and SepPDO (0.72 and 0.66, respectively; Table 5.4). MayPDO and SepPDO were not as highly correlated (0.46) as some other possible combinations. The list of candidate variables (after screening for correlated variables) included SNWTT, AprUP45n, OctUP45n, NovUP45n, MayPDO, and SepPDO. The model selection process was described in methods (p. 115). Text was modified to include the one non-significant result for the current time series. The time period we call current is a short time series, so the result that MayPDO and SNWTT became significant without NovUP45n should not be surprising. The correlation between SeptPDO and SNWTT was also very small.

BPA Comment: Results: Snake River and Downriver SAR Comparisons (pp. 131-136)
• The CSS upstream-downstream comparison of SARs is based on the performance measureU/D =$S_{LGR\text{-}BON}/S_{JD\text{-}BON}$. If there is no differential post-JD mortality for upstream fish, then we expect U/D to equal $S_{LGR\text{-}JD}$, inriver smolt survival from LGR to JD.

…numerous comments continued through…

BPA comment: It is obvious from Table 1 that the value of U/D (and by extension, SARμ) alone does not indicate whether or not "differential mortality" has occurred.

Response: The BPA reviewers seem to misinterpret the purpose of the SAR comparisons, which is to evaluate if the same patterns evident in published SR (spawner-recruit) differential mortality were present in SARs. The purpose was stated in the Chapter 5 introduction (p. 106): "*our specific interest … is whether upriver/downriver differences in SARs for wild and hatchery* *stream-type Chinook were consistent with the differential mortality estimated from SR [spawner-recruit] models for wild populations"* Previous SR analyses indicated there was a systematic increase in mortality for Snake River populations, which did not occur in the downriver populations, associated with the construction and operation of the FCRPS (e.g., Deriso et al. 2001; Schaller et al. 1999).

The reviewers' comments contain a purported comparison of survival from John Day to Bonneville Dam with the ratio of SARs from upriver and downriver stocks (Table 1), and assert that this comparison would be more appropriate than a SAR comparison that indexes smolts leaving the production areas (i.e., at the first dam). There are two problems with the reviewers' approach. Their proposed approach is inconsistent with the original SR definition of differential mortality (e.g, Deriso et al. 2001), where spawners were indexed at the spawning grounds and recruits were indexed at the Columbia River mouth (p. 116, lines 29-31). Second, the reviewers propose to account only for the passage mortality experienced by in-river migrants and not that of transported smolts (the migratory route the majority of fish experience). One could, in theory, fix the smolt indexing location at any number of locations (JDA or BON), but this would be a very different analysis, and not consistent with the SR based estimates of differential mortality. It is not clear what the reviewers' proposed adjustment only for in-river survival would accomplish, other than further confuse this issue.

BPA Comment: Page 132, Table 5.8
• How are the SARs for downriver wild Chinook salmon estimated? If simple return ratio, why not use same method for Snake River fish?

Response: The methods for John Day wild Chinook SARs are described on p. 116-117 and Table 5.7. As explained in Appendix B, Snake River annual SARs required weighting by study category (T_0, C_0, C_1) to reflect their true proportion in the run-at-large. Because John Day smolts were not experimentally separated into different study categories, there was no need to perform this weighting for these fish.

BPA Comment: Page 134
• Lines 2-5: The CSS claims that the SAR to BON is always higher for the downriver (hatchery) fish, but that is not true for 2003.
• Lines 13-16: The reason given for not providing a confidence interval on SAR for downriver fish in 2004 is because an average survival to BON from previous years is used. However, that survival is not known without error, so a measure of uncertainty should be reported on survival to BON for 2004, and that error could be propagated to produce a CI on SARs.
• In general, the CSS addresses uncertainties incorporated by using a single downstream hatchery stock when the upstream/downstream results show no effect of the hydrosystem (i.e., for hatchery Chinook salmon), but not when the upstream/downstream results do imply hydrosystem effects (i.e., for wild Chinook salmon). This sounds like an inconsistent approach.
Page 135, Table 5.10
• In some years, upriver SAR > downriver SAR for hatchery Chinook salmon, despite additional inriver migration for upriver fish. Presumably, this result is unexpected and should be

addressed. Such results may be due to large measurement error that obscures the relationship or the upstream/downstream pairing is a mismatch.

Response: The draft text in question (p. 134 lines 2-5) does not claim the downriver hatchery SAR was "always" higher; we added the word "generally" to avoid misinterpretation in the final version. In the future, a CI for the SAR of downriver fish in 2004 could be generated with this measure of uncertainty. The point of the reviewers' comments about the single hatchery stock is not clear. The downriver wild aggregate is comprised of three populations, and CSS has proposed adding more populations to reduce uncertainty from this factor. We are simply noting on p. 134 that in addition to use of a single downriver hatchery stock, that use of hatchery fish as surrogates of wild fish performance has additional potential confounding factors: hatchery practices, disease, rearing conditions and fitness.

BPA Comment: Results: Comparison of Biological Characteristics of Snake River and downriver smolts (pp. 136-143)
Page 139
• Lines 2-3 It says that there is a significant ($P < 0.001$) difference in density-adjusted mean fork lengths of 106 and 106 mm (for IMNTRP and JDAR1), and separately of 100 and 100 mm (for SALTRP and SNKTRP).
• Lines 6-7: The report is inconsistent when it says 74 mm vs. 121 mm in fork length is not significant, especially considering that they previously defined any differences >5 mm to be biologically significant.

Response: As noted on page 120, because the sample sizes were very large we had the ability to detect small fork length differences with a high degree of statistical significance. We changed the text to more accurately reflect the results of comparing fish sizes for John Day to Snake Basin populations.

BPA Comment: Page 141, lines 11-13
• "Smolts from upriver populations and downriver-origin smolts migrated at a similar rate, once their different migration distances were accounted for." What does this mean? Their migration "rate" (i.e., distance traveled per unit time) already accounts for differing migration distances.

Response: We changed the sentence to say: This comparison demonstrates that smolts from upriver populations and downriver-origin smolts migrated at a similar rate.

BPA Comment: Page 142
• Lines 15-17 – The observation that upriver smolts took longer to travel to BON than downriver smolts is not surprising since they leave at the same time and travel at the same rate, given that upriver smolts have farther to travel.

Response: As discussed above, distance did not change as a result of FCRPS development; water velocity (and WTT) did change. In the impounded river system, smolts are moving at approximately the rate of water velocity (e.g., Fig. 5.22); current average WTT is about 19 days. WTT for Snake River stocks before FCRPS development was only about 2-3 days; Snake River smolts were historically able to arrive at the estuary more in synchrony with their morphological, physiological and behavioral development (e.g., Budy et al. 2002; ISG 1999). If the optimal estuary entry timing for an individual smolt is 12 days after passing LGR, on average, it will arrive a week later than optimal, given the current FCRPS configuration and management.

<u>BPA Comment</u>: Results: SARs by Bonneville Arrival Timing (pp. 144-146)
Page 143, lines 1-2
• The "pattern of delayed arrival" was not consistent across years, as is stated – See years 2000 and 2003.

<u>Response</u>: We added the word "generally" to final draft.

<u>BPA Comment</u>: Page 144
• Lines 11-14 – What reference point is used to determine that upstream smolts experience delayed migration?
• Lines 16-18 – What does "significantly experienced lower SARs" mean? Does this mean that the difference or ratio between the SAR for wild upstream Chinook and wild downstream Chinook was statistically significant? Biologically significant?

<u>Response</u>: The reference point is the large reduction in water velocity from historical conditions discussed above, and strong observed relation between FTT and WTT (see also Chapter 2).

The sentence on line 16-18 was reworded: "All groups of Snake River wild Chinook experienced significantly lower SARs (Bonneville to Bonneville) than John Day wild Chinook within the same arrival time period and for the season, based on non-overlapping 90% CI." This difference in SARs would be statistically significant, and considering that the point estimates differ by about 2-fold, also biologically significant.

<u>BPA Comment</u>: Page 145, Figure 5.23
• Binomial confidence intervals are shown, but error is not binomial for C0, C1, and T0. Recalculate appropriately.
• In some years, large numbers of upriver migrants are omitted from the analysis by restricting attention to 16 April – 31 May window.

<u>Response</u>: The number of smolts arriving at Bonneville is a known quantity; therefore applying binomial confidence intervals is appropriate.

The purpose of this comparison was to compare SARs from the same arrival timing, therefore, because there were so few John Day smolts during the late arrival period it was omitted from the analysis. Note that all the data are available in table 5.16.

<u>BPA Comments</u>: PIT-tag SARs versus SAR of run-at-large (p. 147)
• Lines 3-5: Are run-reconstruction SARs and CSS SARs mathematically comparable? Justify.
• Lines 12-19: Assumptions necessary for the run-reconstruction SARs are discussed, but not assumptions for the CSS SARs.

Response: The run reconstruction SARs in the draft report inadvertently included jacks. This has been corrected to exclude jacks for consistency with the CSS SARs, and text has been modified. Methods and statistical assumptions for the CSS SARs are covered in Appendix B (and elsewhere) in the report, and the issue of a potential negative bias for PIT-tag SARs was addressed in this section and the discussion.

BPA Comment: Discussion (pp. 148-151)
Page 148
• Lines 21-22: The limitations of small sample size cannot be avoided by using multi-year methods, as indicated here. Multi-year methods result in conclusions that are based on many uncertain estimates (due to small yearly sample sizes), instead of based on only a single uncertain estimate. This simply expands the problem of small sample size.

Response: The text referenced refers to the analyses presented in Chapter 4. The Chapter 4 methods explicitly deal with the effects of small sample size. They produce an estimated mean, weighting by sample size, and so account for small sample size, rather than "expanding the problem". Sampling variance is estimated and removed from total variance to get a truer estimate of actual inter-annual variance in SARs, and hence in the ratio of SARs as well.

BPA Comment: • Lines 29-31: WTT is named the "best" predictor variable for SARs, but it is not clear that the CSS considered other inriver covariates.
• Lines 37-38: It was found here that WTT influences the smolt migration rate. But JDAR1 and Snake fish have similar migration rates. Did they have different WTT? This needs to be addressed.

Response: The actual language indicated that SARs were best described by WTT and certain ocean/climate variables. As explained in the model results, selection criteria (AICc and BIC) identified the best models, which always included the WTT variable. We agree with the reviewer that other candidate migration variables should be investigated in the future. Inspection of Figure 5.22 (old Fig. 5.21) on page 143 clearly shows that WTT between the first and third dam experienced by John Day migrants (2-5 days) was shorter than the WTT experienced by Snake River migrants (7-11 days).

BPA Comment:• Lines 42-43: SARs of downriver fish are compared to SARs from upriver fish, but these SARs are estimated over DIFFERENT reaches and distances, so we expect them to be different. The CSS needs to investigate whether the differences are more than expected.

Response: We addressed this issue above. Briefly, spawner-recruit (SR) differential mortality estimates (1.1 – 1.5) suggest about a 3-4 fold ($e^{-1.1}$ to $e^{-1.5}$) difference in life cycle survival after completion of the FCRPS. Migration distance did not change after FCRPS development; therefore, it is hard to see how different distance would drive the differential mortality response in SR. Our primary interest was whether SARs indicate the same differential mortality as was evident from the SR analyses during the post-dam period. For wild upriver/downriver SAR contrasts to date, we see a similar level of differential mortality as was evident from previous SR analyses.

BPA Comment: Page 149
• Lines 39-40: "Hydrosystem migration rates did not differ between groups but were strongly influenced by water travel time." It is not clear how to interpret this statement. Did groups have different water travel times but the same migration rate? Or did they have the same WTT? Or was migration rate and travel time examined on an individual fish basis, instead of a group basis? It is not clear.
• Lines 41-46: Distance to travel is not considered as a factor of travel time.
Page 150
• lines 1-6 – It is claimed that the "potential confounding effects due to life history differences are probably negligible," but the CSS does not attempt to model SAR using both the upstream/downstream designation and the life history differences. Additionally, the effect of distance to travel was ignored. A model that includes all possible factors affecting SAR should be considered, in order to claim that it is the hydrosystem rather than other factors that cause the difference in return rates.

Response: Sentence in question was modified to: "When Water Travel Time was incorporated in the analysis, there was no difference in migration rates between groups." The issue about distance was addressed above.

BPA Comment: In general for Chapter 5
• In order to determine if there is a biological difference that explains any differences in SAR between upriver and downriver stocks, model SAR using fork length, migration date, arrival timing, year, in addition to upstream/downstream classification. Is upstream/downstream effect significant, given presence of all others?
• Looking at population differences in fork length, migration date, etc., one at a time, is reasonable for initial data exploration, but insufficient for conclusions about the significance of the upstream/downstream effect.

Response: As discussed above, the estuary arrival timing distribution for Snake River juveniles is largely a response to the FCRPS (delay of in-river migrants, combined with a mix of project delay and barging for the transported individuals), and may not be an appropriate "independent" variable. We could pursue the remainder of the suggested analysis in future reports. However we note that SARs have been about 4-fold higher for the downriver wild populations, and none of the biological characteristics examined to date exhibit differences that would provide a plausible alternative explanation for this level of differential mortality.

BPA Comment: Throughout Chapter 5
• Typos are made in references to tables and figures throughout the entirety of Chapter 5.
• Pages 139–144: The reader is referred to a nonexistent figure for release site abbreviations.

Response: Addressed.

Chapter 6

BPA Comment: Page 154, line 9
• The notation RY has not been defined. The context suggests Return Year, but Release Year is also a possibility.

Response: MY (migration year) and RY (return year) have been defined in the final draft.

BPA Comment: Page 154, Tables 6.1 and 6.2
• Pooling migration success data across migration year and return year is valid only if those factors are nonsignificant. Perform test of homogeneity.
• Also applies to Page 155 (lines 17-23); Page 156 (Table 6.3).
Page 155, line 41
• Was return year modeled as a fixed or random effect? Most blocking factors are modeled as random effects, although there are times when a fixed effect is more appropriate.
Page 156, Table 6.3
• Chi-squared tests indicate whether there is a difference in perceived upriver adult survival across juvenile migration groups, but they do not indicate the nature of the difference. The p-values reported do not indicate that the actual ranking in the Success Rate Ranking column is significant, simply that at least one of the juvenile migration groups had a significantly different success (survival) rate than the others. One-sided tests should be performed comparing pairs of juvenile migration groups in order to test the significance of the ranking

Response: The first three bullets suggest our presentation of survival and travel time analyses (each being separate efforts) may have been somewhat confusing. As we analyzed success on a year-by-year (i.e., migration or return) and pooled (and hatchery-specific) basis using separate χ^2 tests, there was no explicit model structure for this exercise. Given pooled, MY-, and RY-specific test results, however, a formal test for year effects (a factor of secondary interest) will not change our conclusions about the principal factor of interest (outmigration experience). This is especially true given the results from our logistic regression analysis. In contrast, our GLM-based analysis of travel-time data did incorporate an explicitly defined model structure; in this exercise, return year was modeled as a fixed effect (Bullet 3).

Regarding the reviewer's last comment (Bullet 4), we presented the rankings in Table 6.3 to emphasize the consistency of ranking patterns across tests and groups. While the reviewer is correct that *post-hoc* one-sided tests could more finely resolve where the lack of homogeneity exists in the data in a purely statistical sense, this does not necessarily preclude discussion of general patterns.

BPA Comment: Page 157, Figure 6.1
• Needs error bars or confidence intervals.

Response: The estimates and CI are in Appendix D (Tables D-32 – D-36); showing the CI on the figure would result in a very cluttered graphic.

BPA Comment: Page 158, Figure 6.2
• Needs error bars or confidence intervals.

• The interpretation of Fig. 6.2, showing the proportion of LGR-detected adults and jacks detected at hatcheries, depends on the detection effort at each hatchery in each year. Without that information, it is useful only for comparing transported to inriver fish. It appears that transported fish had slightly better survival from LGR to the hatcheries, but without error bars and without information about detection effort (and harvest pressures, etc.), no real conclusion can be reached from Fig. 6.2.

Response: The 90% CI were added to figure 6.2 in the final draft. Transport and in-river CIs overlap for all years, indicating little evidence of a difference in detection probability at the hatcheries.

BPA Comment: Page 159, lines 11-15
• The overall average perceived BON-LGR adult survivals for the three migration groups are not very useful without standard errors or confidence intervals.
• It is not clear how these average survivals were computed. Were yearly estimates weighted by the number of fish returning in each year? Or were migration year estimates averaged?
• Given the finding that return year is a significant factor in perceived upriver adult survival (from the logistic regressions presented later in this chapter), pooling data over return year is not warranted.

Response: The summary that the BPA reviewer states is not useful without a presentation of confidence intervals is inaccurate, as we do present 95% confidence intervals graphically (Figure 6.4). The average "success" proportions (equivalent to the reviewer's 'perceived survival') reported on Page 159 (and plotted with 95% CIs in Figure 6.4) were computed using the pooled data (i.e., the 'Combined' field) in Table 6.1. Thus, the values presented in the figure and reported in text are unweighted averages. We also computed weighted (by return or migration year sample sizes) estimates, however, and they are virtually identical: weighted averages for Hatchery In-river, LGR, and LGS groups were 0.83, 0.76, and 0.81, unweighted values were 0.84, 0.77, and 0.81, respectively); weighted averages for Wild In-river, LGR, and LGS groups were 0.87, 0.74, and 0.90, unweighted values were 0.87, 0.76, and 0.89, respectively). Reviewers' last bullet appears incorrect. RY was significant in the travel time test, but was not in the logistic regression for adult survival.

BPA Comment: Page 160
• Lines 16-19: The model evidence ratio does not indicate that one model is "more likely" than another, in a Bayesian sense. Rather, it means that there is more evidence for one model compared to the others.
− Also applies to results for wild Chinook salmon (p. 161, lines 22-23).
− The highest evidence ratio for the best model for wild Chinook salmon (p. 161, lines 21-25; p. 162, Table 6.6) is at most 4, thus there is not clear evidence that transportation is an important factor in determining adult migration success when compared to environmental factors.
• Lines 29, 32: It is not clear how the confidence intervals on the odds ratios are computed. Provide explanation. Asymptotic normal-theory confidence intervals are considerably narrower than those reported, and do not include 1 for either LGR-transport fish or LGSdown fish. If the

confidence intervals were based on a t-distribution, the degrees of freedom should be reported (Table 6.5).

Response: The reviewer's first comments are a matter of semantics, not a technical or analytical issue necessarily. We used these model fit criteria as one (among others provided) to judge which model(s) best explained the observed data. The reviewer is taking literary license with what we said in the text, as our conclusion based on model results was that there is stronger support for a transportation-legacy hypothesis than an environmental conditions-only hypothesis. The confidence intervals shown on page 160 for the odds ratio of parameter **LGR** relative to parameter **In-river** and parameter **LGS down** relative to parameter **In-river** are obtained from running a binomial logit in SYSTAT (logistic regression) with a categorical variable **transport** (split on three levels: In-river = 0, LGR = 1, and LGSdown = 2) and the other non-categorical variables modeled. An exponential transformation of the logistic regression parameter estimates for LGR and LGSdown will provide the odds ratio of these parameters relative to In-river. SYSTAT prints out the odds ratios and 95% confidence intervals directly, but they may also be obtained by taking exponents of the logistic regression parameter estimates of LGR and LGSdown and their 95% confidence intervals. Table 6.5 shows a logistic regression parameter estimate and standard error for LGR of -0.446 and 0.092, respectively, from which the 95% CI is $-0.446 \pm 1.96 \cdot 0.092$ ➔ (-0.6263, -0.2657). The exponential transformation results in a odds ratio of 0.64 and 95% confidence interval of (0.53, 0.77) as shown on page 160. Likewise, Table 6.5 shows a parameter estimate and standard error for LGSdown of -0.212 and 0.123, respectively, from which the 95% CI is $-0.212 \pm 1.96 \cdot 0.123$ ➔ (-0.4531, 0.0291). The exponential transformation results in an odds ratio of 0.81 and 95% confidence interval of (0.64, 1.03). It is not clear how the BPA reviewer computed narrower asymptotic normal confidence intervals for the odds ratio.

BPA Comment: Page 161
• Table 6.5
− Degrees of freedom should be reported for each parameter estimate.
− Surprisingly, warmer temperatures were associated with higher perceived adult survival. Perhaps temperature is confounded with run (spring versus summer).
• Lines 29-30: The odds ratio is misinterpreted here. An odds ratio of 0.5 does not mean that the probability of success of LGR-transport fish is half that of inriver fish. If the probability of success (i.e., perceived adult survival from BON to LGR) is for LGR-transport fish, and is for inriver fish, then:

[Odds ratio = ½, then $P_{LGR} = P_{inriver}/(2-P_{inriver})$]

This means that the probability of success of LGR-transport fish depends on the value of the success probability for inriver fish, as demonstrated in Table 2 below. Table 2 indicates that for an odds ratio of 0.5, the probability of success of LGR-transport fish is generally greater than half that of inriver fish, except for very small inriver success probabilities, which are not applicable here.

Response: We changed the language on p. 161, lines 28-30 to more accurately reflect interpretation of the odds ratio as follows: "Further, the odds ratio estimate for the LGR group (estimate: 0.46; 95% CI: 0.26-0.84) indicates that these adults had significantly lower odds of surviving their BON-LGR migration than in-river outmigrants (i.e., the 95% CI did not include 1)."

BPA Comment: Pages 162-163, Hatchery Chinook arrival and travel time ANOVAs
• For both arrival time and travel time, the interaction term between return year and juvenile migration (outmigration) method was significant. This affects interpretation of the main effects of both return year and outmigration method, so conclusions based on the main effects alone are invalid.
• The ANOVAs should be included in the report.

Response: The reviewers are mistaken in implying we drew conclusions about main effects on arrival time and travel time. We accurately reported the results of the interactions.

BPA Comment: Page 164, lines 38-40
• How much of TIR or D is explained by observed differences in perceived upriver adult survival between inriver and LGR-transport fish?

Response: In the conclusions, we were simply noting that a portion of deviation in *TIR* and *D* may be attributable to survival differences occurring in the mainstem after adults return .We did not attempt to quantify this phenomenon. Based on future priorities, this could be a focus for future studies.

Chapter 7

Page 168 (lines 31 and 46) – two comments regarding input values to the simulator program.

> i). Survival from release to LGR = 0.95 seems high, and does not correspond to year 2000 data used as basis for default values.
>
> Response: Migration year 2000 data was used to establish the default curves for survival rates, collection efficiency at dams, inter-dam travel times, and the initial arrival timing distribution at LGR. The survival from release to LGR was simply set at 95% to reflect a typical survival rate from the head of the hydrosystem at Lewiston to LGR. The release size of 32,000 fish was aimed at providing an arrival population at LGR of approximately 30,000 fish, which is in the range observed with wild Chinook as well as hatchery Chinook from Rapid River and McCall hatcheries in several years. Since program computing time increases somewhat exponentially as release number increased, a higher release number and lower survival rate from release to LGR to achieve approximately 30,000 fish arriving LGR would have increased the overall computing time without affecting the simulation outcome.
>
> ii). An SAR=0.03 seems high, given that observed SAR has been lower than the target value of 2% in most years according to this report.

Response: Only one SAR level was simulated for this report, with those fish having capture histories reflective of a particular study category getting assigned an adult return based on the random binomial draws from the number of smolts in that particular study category prior to any expansion to LGR equivalents. Since the assumptions being tested in the 12 scenarios run to date related to temporal changes in inter-dam reach survival rates and collection efficiencies at dams, and not to temporal changes in SARs based on timing of smolt arrival at LGR, we did not need to run more than one level of SAR. Had we set the input SAR at 2% or 1%, our ability to investigate biases caused by violation of the CJS model assumption that "all fish in a release group have equal detection and survival probabilities within the same river reach or at the same dam" (Assumption #2 in Appendix C) would not have been affected. The resulting population variability about the SARs, *TIR*s, and *D* would increase as the input SARs level got smaller, but this effect is unrelated to the CJS model assumption being tested.

Page 169 (line 9) BPA Comment: comment that the joint probability of survival from BON to TWX and detection at TWX =0.10 is high based on past years.

Response: A lower joint probability could have been used, but it would not have affected our evaluations of impacts of violations of Assumption #2 described in the previous response. We allowed temporal changes to occur in reaches and dams between LGR and MCN, and maintained the same default inputs for all reaches and dams below MCN as well as at the trawl in the 12 scenarios tested. A lower joint probability assigned at the TWX would have reduced the number of smolts caught in the trawl and thus increased the population variability for the S_R and D parameters to some extent, but again as in the previous response, this effect is unrelated to the CJS model assumption being tested.

Page 170 (lines 9-11, 21-23, and 39-41) and page 171 (lines 8-10) – comment that survival probabilities used in simulation scenarios #5, 7, 10, and 12 include inriver survival probabilities >1, when the variable day is 0 or very low. Inriver survival should be parameterized using only admissible parameter values (i.e., ≤1) and included in this report.

Response: In the simulator program, we have constraints on the daily values taken from the parabolas and linear trends to avoid the problems the reviewer expressed. The survival rate and collection probabilities are not allowed to exceed 0.95 or drop below 0.05, in order to keep the random beta distribution draws from occasionally trying to return an undefined value (>1 or <0), which terminates the run. Figures 7.8 and 7.9 show the how this constraint changes the steepest linear trends evaluated to flat lines before or after certain dates. In the methods section where the trend lines for the 12 simulation scenarios are presented, we will add text to indicate that daily values taken from the parabolas or linear trends are constrained between 0.05 and 0.95 prior to these values being used in beta draws for survival rate and collection probabilities that are finally used in the binomial draws for numbers of fish surviving as well as collected each day within the various inter-dam reaches and dams.

Page 171 (last paragraph) and page 174 – comment regarding SIM-2 where the emphasis on the T and R groups is confusing. The review comment goes on to suggest that a simpler method of assessing the effect of detection-influenced survival would be to simulate date under the *scenarios described (post-turbine survival < post-bypass survival < post-spill survival, with varying proportions of undetected fish passing via turbine or spill) and examine estimates of C_0, C_1, and T_0.*

Response: The simulator program was designed to address the impacts on CJS estimates of survival rates and collection probabilities when the underlying "true" survival rates and collection probabilities are changing temporally. This condition causes violation of the CJS model assumption that "all fish in a release group have equal detection and survival probabilities within the same river reach or at the same dam." It was not designed to address the impacts caused when prior detection history causes a change in later downstream survival rates and collection efficiencies. With the start in 2006 of pre-assigning PIT-tagged fish into Group T which reflects the experience of the run-at-large (untagged and tagged) and Group R which is used for estimating in-river survival rates. The attempt in the draft report to address the potential impacts of prior detection history in the indirect approach utilized was determined by the CSS Oversight Committee to be too ambitious given the tight deadlines for the 10-year report. Therefore, we have deleted SIM-2 from the report.

Page 175 (last paragraph) – comment states that it is not clear if the "true" survival parameters used to compute LGR equivalents are averages of seasonal survival parameters, or if LGR equivalents are computed on a daily basis and then summed over the season. Give the temporal variation in survival parameters introduced in these simulations, the latter approach should give a better representation of the "true" C_0, C_1, and T_0 groups. Clarify approach and, if necessary, rerun simulations.

Response: The known (i.e., "true") S_2 and S_3 used to convert smolt counts to LGR equivalents are obtained in three steps: 1) survival rate from LGR to LGS is obtained by dividing the LGR computed "known" number of fish remaining inriver (after subtracting off the removals for transportation) for the season into the LGS "known" number of fish surviving there, which is computed by summing over the season the daily number of fish assigned as survivors based on binomial draws each day with survival rates obtained from the daily trend relation (parabola or linear); 2) the travel time from LGR to LGS distributions will shift the surviving fish at LGS into their starting dates there; 3) survival rate from LGS to LMN is obtained in the same manner as step 1 (simply substitute LGS for LGR and LMN for LGS). When step 3 is completed the "known" number of fish surviving to LMN is obtained. This process has produced S_2 and S_3 that are based on total "known" fish arriving the downstream dam divided by total "known" fish continuing inriver from the upstream dam. This process produces the proper "known" parameters S_2 and S_3 for use in converting downstream smolt counts into LGR equivalents. The approach preferred by the reviewer would be much more difficult to implement, but should provide the same starting population at LGR, if done correctly.

Page 194 – comment is split into two parts due to length followed by answers.
Comment – The CSS uses results of the second set of simulations to address how to best analyze data using the NPT approach, in which tagged fish are pre-assigned in to migration groups: T (transport) fish are transported upon their first detection at a transport dam; R (river) fish are returned to river upon all detections. Using the $C_0/C_1/T_0$ approach to analyze data with pre-assigned migration groups is not intuitive.

Response: This is incorrect as stated. The goal of the second set of simulations (now dropped as too preliminary) was not to address how to best analyze data using the NPT approach, which has been implemented in the CSS starting with the 2006 migration year. Rather it was aimed at showing how the categories C_0, C_1, and T_0 utilized in CSS analyses may best be computed. In the current CSS, annual estimates of overall SAR must be computed as a weighted combination the category-specific SARs, where the weights are the proportion of the run-at-large (untagged and tagged) represented by each category. Using the NPT approach, Group T will provided the annual estimate of overall SAR directly. The reviewer also mischaracterized the fish in Group T as being transported upon their first detection at a transport dam. At collector dams, Group T fish go the direction of the untagged fish, regardless of whether that is to raceways for transport or back to the river. Likewise, if untagged fish are being transported from a dam, then any fish in Group T detected at that dam will also be transported, regardless of whether that fish had been previously detected at dam upstream.

Comment: It would be simpler and more defensible to simply compare the SAR of the T group to the SAR of the R group. All "R" fish will have migrated wholly inriver, while some "T" fish will have been transported and others (undetected) will have migrated inriver. The comparison of SAR(T) to SAR(R) is more easily interpreted for management, because the alternative to transportation is to return detected fish to the river, whereas the transportation alternative being tested in the SAR(T0) vs. SAR(C0) comparison is not clear.

Response: When analyzing the data collected from migration years 2006 and later, we will be comparing SAR(T) to SAR(R) as the reviewer suggests, but this does not preclude the utility of additional comparisons among all three study categories C_0, C_1, and T_0. Just as we have a time series of $SAR(T_0)$, $SAR(C_0)$, and $SAR(C_1)$, and overall $SAR_{LGR\text{-}to\text{-}LGR}$ (akin to SAR(T)), we will also be able to compare SAR(R) data with prior years by substituting $SAR(C_1)$ for $SAR(T_0)$ in the formula of overall $SAR_{LGR\text{-}to\text{-}LGR}$ for pre-2006 migration years. The reviewer failed to include the fact that in addition to fish transported and those undetected at collector dams, Group T may include fish bypassed at Snake River collector dams during April and early May under the policy begun in 2006 of delaying the start of transportation at those dams.

Chapter 8

Page 198 Lines 35-38: The trend of performance measures for wild fish mimicked the overall trend of performance measures for the collection of hatcheries, but did not agree well with the trend from any single hatchery across all years. It is not clear which single hatchery could be used to make inference to wild fish. Also applies to Pages 199-200, bullet (b) of Chapter 5 summary.

Response: Hatchery Chinook salmon and wild Chinook salmon responded nearly identically to environmental and/or seasonal conditions in terms of their fish travel time, instantaneous mortality rates, and survival rates in the LGR-MCN reach. Thus, hatchery Chinook salmon provide valuable information on the response of wild Chinook salmon to conditions experienced in the hydrosystem.

Differential mortality between upriver and downriver stream-type Chinook populations has been estimated for wild populations from both spawner-recruit (Schaller et al. 1999; Deriso et al. 2001; Schaller and Petrosky 2007) and PIT-tag SAR (CSS study) data sources. The CSS also investigated whether a similar level of differential mortality was present between PIT-tag SARs for five upriver and one downriver hatchery Chinook populations. Because biological characteristics of a population could differentially influence survival to adult return (see above), we also summarized hatchery pre-smolt FL at the time of tagging, and hatchery smolt arrival timing distributions entering the hydrosystem (LGR or BON) and arriving at the estuary (BON).

Upriver and downriver hatchery spring/summer Chinook SARs did not show the same level of differential mortality as was apparent from the wild populations. Survival of hatchery fish is subject to additional fitness and rearing factors that may not affect wild populations. CSS currently has the ability to compare SARs from a single downriver hatchery (Carson NFH) with those from five Snake River hatcheries. Additional candidate populations relevant to these SAR comparisons from downriver hatcheries of the Interior Columbia include Klickitat, Warm Springs, and Round Butte (depending on fish health constraints). Future monitoring should also consider incorporating PIT-tag SARs from the upper Columbia region to expand these regional comparisons.

Although Snake River hatchery Chinook exhibited a generally more positive response to transportation and relatively lower levels of differential mortality than wild populations, annual SARs of wild and hatchery Snake River Chinook were highly correlated. In view of this high correlation, continuing the CSS time series of hatchery SARs will be important to augment wild Chinook SAR information following future years of low escapements, in addition to providing valuable management information for the specific hatcheries. One advantage of the CSS study is that tagging takes place at the hatcheries and in the tributaries for wild populations. This approach allows for detecting different responses to management actions for different components of the wild and hatchery aggregate groups, unlike approaches that only tag at the upper most dam. Finally, it is of interest to the region of how the specific hatchery groups respond to the hydrosystem management actions. The reviewers suggest a much smaller number of PIT-tagged hatchery fish could be used. We believe that the sample sizes should be periodically reviewed based on updated survival estimates, and regional monitoring and evaluation needs.

Page 198, Report confidence intervals for results (e.g., geometric means).

Response: Confidence interval results are presented in Appendix D.

Page 199 • Lines 32-35: The inference made from declining SAR(C_1) over the season to hydrosystem-caused post-Bonneville mortality is unfounded. There are alternative possible causes of post-Bonneville mortality, including temperature, pollution, disease, and seasonal changes in estuary conditions. No conclusions about the relative importance of the various potential sources of mortality can be reached here.

• Lines 40-42: The CSS claims that Snake River wild steelhead SARs averaged less than 2%. It is difficult to confirm this statement, because the annual SARs are not presented in tabular form in this report. However, Fig. 3.12 suggests that average annual SAR for wild steelhead may be greater than 2%. Document annual SARs in the table and explain apparent inconsistency.

Response: No unambiguous demonstration of the effect is claimed; the report states that the declining SAR is "consistent with the hypothesis" of protracted migration-induced mortality. Most of the commenter's listed alternative causes are actually mechanisms which could cause mortality due to protracted migration. For instance, temperatures increase over the season (for spring migrants). Disease expression can be affected by protracted migration, through delaying of saltwater entry relative to smoltification and through exposure to higher temperatures. Seasonal changes in the estuary are another likely candidate for mortality induced by late arrival of smolts. If the commenter has evidence that seasonal distribution of pollutants in the estuary can explain such a dramatic drop in post-Bonneville survival over the season, we would be eager to see it.

Evidence for wild steelhead SARs averaging less than 2% can be found in Table D-19, where transport SAR averages slightly over 2%, but in-river SARs average less than 1%. Annual overall steelhead and Chinook SARs are also found in Appendix E, which will be included in the next draft of the report. Further, the Chapter 4 weighted mean wild steelhead overall SAR is 1.95% (Figure 4.4).

Page 200, Lines 8-14: The CSS did not compare the ratio of upstream and downstream SARs to in-river survival between Lower Granite and John Day, so the conclusion that upstream fish experience extra mortality caused by the hydrosystem is unjustified. Also applies to Page 200 (lines 33-34).

Response: The reviewers' comments contain a purported comparison of survival from John Day to Bonneville Dam with the ratio of SARs from upriver and downriver stocks (Table 1), and assert that this comparison would be more appropriate than a SAR comparison that indexes smolts leaving the production areas (i.e., at the first dam). There are two problems with the reviewers' approach. Their proposed approach is inconsistent with the original SR definition of differential mortality (e.g., Deriso et al. 2001), where spawners were indexed at the spawning grounds and recruits were indexed at the Columbia River mouth (p. 116, lines 29-31). Second, the reviewers propose to account only for the passage mortality experienced by in-river migrants and not that of transported smolts (the migratory route the majority of fish experience). One could, in theory, fix the smolt indexing location at any number of locations (JDA or BON), but this would be a very different analysis, and not consistent with the SR based estimates of differential mortality. It is not clear what the reviewers' proposed adjustment only for in-river survival would accomplish, other than further confuse this issue.

Page 200, Lines 23-24: The claim is made that that the CSS shows clear evidence of delayed estuary entry of Snake River in-river smolts, caused by passage through the hydrosystem, on the basis of comparisons with John Day smolts. This is not true. The CSS found that Snake River and John Day smolts (1) initiate migration at the same times, and (2) migrate at similar rates through the first three dams passed. Given the extra distance traveled by the Snake River smolts, *it is not surprising that Snake River smolts enter the estuary later than John Day smolts. The CSS analysis would be more useful if it had compared the observed and expected arrival dates of the Snake River fish, given their migration initiation date, migration rate (through the first three dams), and distance to travel.*

Response: The BPA reviewers appear to be confused on the purpose of the upriver/downriver analysis, which was stated in the Chapter 5 introduction (p. 106): "*our specific interest … is whether upriver/downriver differences in SARs for wild and hatchery stream-type Chinook were consistent with the differential mortality estimated from SR [spawner-recruit] models for wild populations"* Previous published SR analyses indicated there was a systematic increase in mortality for Snake River populations, which did not occur in the downriver populations, associated with the construction and operation of the FCRPS (e.g., Deriso et al. 2001; Schaller et al. 1999). In the SR model formulations, any differences in smolt mortality caused by different travel distances would be incorporated into the intrinsic productivity (Ricker "a"). Obviously, the migration distance for upriver and downriver populations did not change over the time period of FCRPS development; Water Travel Time (WTT), Fish Travel Time and hydro impacts did change with this development. WTT for Snake River stocks before FCRPS development were only about 2-3 days; Snake River smolts were historically able to arrive at the estuary more in synchrony with their morphological, physiological and behavioral development (e.g., Budy et al. 2002; ISG 1999). Available evidence from a mostly free-flowing migration corridor (Whitebird trap on the Salmon River to Ice Harbor Dam) also suggests smolt survival was high before FCRPS development (Raymond 1979). Applying the survival per mile from the Raymond study the information suggests that the historic survival from Lewiston to Bonneville dam was over 90%.

Page 200, Lines 26-30: The conclusion that differing seasonal SARs for upstream versus downstream smolts is evidence of delayed mortality ignores possible alternative explanations, including potentially different ocean residencies.

Response: Based on the weight of evidence in the peer-reviewed literature, it is apparent that the highest level of mortality takes place in the first year of ocean residence.

Page 200, Lines 37-42: It appears here that wild and hatchery Chinook salmon transported from LGR always had 10% lower SAR than fish passing through the hydrosystem by alternative routes. It should be noted that the effect for hatchery fish (4% to 7%) was considerably less than the effect for wild fish (15%), so the 10% effect reported is somewhat misleading.

Response: In this comment, the reviewer has confused adult upstream survival rates with SARs.

Page 202, Lines 11-16; lines 39-41: The claim is made that the CSS addresses the question of whether smolt transportation compensates for effects of the Federal Columbia Power System (FCRPS) on survival of Snake River Chinook and steelhead. This claim extrapolates past the available data. The CSS compares the SAR of transport fish to the SAR of fish migrating in-river. While the in-river fish experience effects of migrating through the FCRPS, available data do not *indicate the magnitude of those effects; this would require comparing the SAR of fish migrating through the FCRPS to the SAR of fish migrating through the same reaches but not through the FCRPS. That is not possible. At most, the comparison of the SAR of transport fish to the SAR of in-river fish indicates whether transportation is a viable management option; it is not equivalent to comparing transportation to migration through the unimpounded river. It is worth noting that the SAR from BON to BOA for hatchery Chinook salmon from the John Day river was less than 2% for 2001 through 2004 (Table 5.10). Regardless of the validity of upstream-downstream comparisons, these low SARs for John Day fish suggest that the hydrosystem is not the only factor in below-target SARs.*

Response: The reviewer has misconstrued the analyses conducted within the CSS. The CSS has monitored the effectiveness of transportation versus in-river migration in the presence of the FCRPS. We have also evaluated those SARs relative to the NPCC's 2-6% SAR objectives. We make no statements regarding survival in an unimpounded river. The reviewer makes references to hatchery Chinook salmon from the John Day River, which do not exist. It is important to note that the wild Chinook SAR from the John Day River has met the NPCC SAR objectives, providing evidence that stocks which migrate through fewer dams can meet these interim survival objectives.

Page 202, 3rd paragraph

Response: The geometric mean is a better measure than arithmetic mean of central tendency for right skewed (log-normally distributed) distributions such as *TIR* and *D*. They both represent ratios of survival rates, for which the ordering (i.e. which is numerator and which denominator) is arbitrary. From Zar (1984, p. 24): "[The geometric mean] finds use in averaging ratios where it is desired to give each ratio equal weight".

The wording about steelhead *D* will be changed to indicate the evidence about whether *D* is in general less than 1 for wild steelhead is ambiguous. The implications of *D* being less than one while *TIR* is greater than one will be noted. The question of whether or not to transport depends in large part on what the alternative to transportation is. The value of *TIR* serves to answer this question in some contexts, but not in others. If the only alternative is simply to allow migration in-river under current configuration and operation, *TIR* is a useful metric. If the range of alternatives included strategies to significantly improve in-river migration conditions, up to and including dam breaching, then *D* tells us more about any expected benefits that might be derived from these alternative strategies.

Page 202, Last paragraph: The CSS compares observed SAR estimates from PIT-tagged fish to the NPCC objectives for SAR (2% minimum, 4% average), without addressing the NOAA finding that PIT-tagged fish have lower survival than untagged fish (as requested by the ISAB). Without knowing the size of the PIT-tag bias, comparisons of PIT-tag SAR to target values are not completely useful.

Response: The introduction to Chapter 5 (p. 105) cites the ISAB (2006) issue that more attention should be given to whether PIT-tagged fish survive as well as untagged fish. Chapter 5 contains a section (p. 147) titled: "Do PIT-tag SARs represent SARs of the run-at-large?" with further discussion on p. 150-151. We agree with the ISAB (2006) conclusion that more attention should be given by CSS *and the Region as a whole* (emphasis added) to the discrepancy of SARs between PIT-tagged and untagged fish. However, the extremely tight reporting requirements did not allow for an examination of all the assumptions and data adjustments currently necessary to estimate SARs of the untagged component. Because the issue involves potential bias of both run-reconstruction and PIT-tag methodologies, resolution will require a collaborative effort among several technical groups in addition to the CSS project.

Contrary to the NWFSC comment that no caveat exists that PIT-tagged SARs may have a bias relative to the NPCC goal, the draft report explicitly stated (p. 147) "[t]he primary concern of negative bias from PIT-tag SARs would be in evaluating whether SARs are meeting NPCC biological objectives (2% minimum, 4% average)." Also, "[i]mplications of bias (if present) would be negligible for relative comparisons of the CSS PIT-tag SAR data, such as between Snake River migrants with different hydrosystem experiences, or between Snake River and downriver populations." We also point to future monitoring and evaluation tasks to help resolve this issue in the future. We note that the 2 to 4 % goal itself was based on analyzes involving tagged fish that presumably experienced some handling mortality relative to the unmarked population.

Page 203, 3rd paragraph

Response: We agree that we have not performed a "comprehensive" analysis of strategies for varying transportation over the season, and we don't believe we implied that. The CSS was not designed primarily for that purpose. However, we have explored seasonal variation in reach survival and transport and in-river SARs and found some interesting results, and we believe that "[Results] have the potential to inform management on when to initiate transportation" is cautiously and appropriately worded.

The C_1 group is the appropriate group of interest for comparison to transported fish for some management questions, and we used this group in the seasonally varying SAR estimates. For instance, if the question is simply "if a fish is collected, then given when it is collected, should it be transported?", this group is appropriate. However, the question of when to turn on or off transportation says nothing about the alternative to transportation, i.e., how the river would be managed for spring migrants in the absence of transportation. Depending on management actions (e.g., high spill at collector projects), there could be a large percentage of C_0 fish in many years. Then, the question is, "When is transport SAR greater than in-river SAR, given that in-river fish would be some mix of C_0 and C_1 fish?" The appropriate weighting of the two in-river SARs would depend on the proportions in each group expected under the particular management regime.

Appendix B

Page B-3, Figure 1 – BPA Comments:

- *The estimators of $\acute{Ø}_1$, $\acute{Ø}_2$, and $\acute{Ø}_3$ are correct.*

- *The figure is somewhat cryptic. The parameters $\acute{Ø}_i$ are not defined, nor are the statistics R_i, $R'_{1\cdot2}$, $R'_{12\cdot3}$, …, r_i, m_i. The reduced m-matrix is not so standard that the CSS should expect all readers to recognize and understand it without further explanation. Provide more detail.*

Response: In order to help the reader understand the notation in Figure 1, we expanded the text to include a detailed description of all notation and concepts being illustrated in Figure 1. The reduced *m*-matrix (detailed in the Burnham et al (1987) monograph) is a useful summarization of all data required to estimate the parameters of inter-dam reach survival rates ($\acute{Ø}_i$) and dam collection probabilities (p_i). It should be familiar to those who have used the CJS model. For those unfamiliar with the CJS model, the schematic with legend should help them better understand the estimation process.

Page B-4 – BPA Comments:

- *The CSS explains that they allow individual reach survival estimates exceeding 100% when computing an overall multi-reach survival estimates. Why, then, do they not allow $S_{JDA\text{-}BON} > 1$ for 2004 for Carson NFH Chinook in Chapter 5?*

Response: In Chapter 5, the CJS based estimate of survival from release at the hatchery to Bonneville Dam was >1 for the Carson NFH Chinook in 2004, not a survival between JDA and BON as stated by the reviewer. In that situation, we felt an average release-to-BON survival rate of the prior years would be better estimate than simply constraining the estimate to 1. This was the first occurrence of a release-to-first dam estimate of survival exceeding 1. Between adjacent reaches, the CJS estimates of survival have an inherent negative correlation, since the estimated population in the tailrace the upper reach becomes the starting population in the next reach downstream. When one estimate is high, the next will be low, and visa versa, as one travels down through all reaches. Therefore, when we take the product of a series of reach estimates to obtain a longer multi-reach survival rate, the reach-to-reach variation is dampened in these longer reaches, thus balancing the effect of some individual estimates being >1. A greater concern is having individual reach estimates of very poor precision lower in the hydrosystem due few fish there. Therefore, we would not used an estimate with CV >25%, and would extrapolate the survival of that reach based on a per-mile survival rate based on the available upriver multi-reach survival rate estimate.

- *The verbal description of the weighted average of survival estimates provided in the second full paragraph is insufficient. An equation demonstrating precisely how the overall survival estimates was estimated is required.*

Response: The distribution of PIT-tagged fish detected at LGR is partitioned into strata. The program allows strata defined by equal proportion of fish per strata or equal number of days per strata. The CJS is run separately fish in each strata and then common reaches across strata are weighted by inverse relative variance times proportion of run-at-large (untagged and tagged fish) for wild Chinook and simply inverse relative variance for Chinook from each hatchery. This approach was only used on Chinook in the early years of the CSS, prior to any analyses on steelhead. Details of the computation of the weighted average survival rate in the j^{th} reach are as follows and has been added to Appendix B:

1. Let A_k = proportion of annual passage index data from Smolt Monitoring Program in the k^{th} stratum
2. Let B_k = theoretical variance of CJS estimates for k^{th} stratum,
   where $B_k = S_{jk}^2$ [$1/r_j - 1/R_j$ + additional terms shown for var($\acute{Ø}_j$) on Page 115 of Burnham et al. (1987)]
3. Let S_{jk} = estimated survival rate of j^{th} reach in k^{th} stratum
4. Weight for wild Chinook is $W_{1k} = (A_k)(S_{jk}^2/B_k)$ in k^{th} stratum
5. Weight for hatchery Chinook is $W_{2k} = (S_{jk}^2/B_k)$ in k^{th} stratum
6. Weighted estimate across k strata for jth reach is:
   $\Sigma\ (W_{1k})(S_{jk})\ /\ \Sigma\ (W_{1k})$ for wild Chinook
   $\Sigma\ (W_{2k})(S_{jk})\ /\ \Sigma\ (W_{2k})$ for wild Chinook

- *The CSS used weights equal to the inverse relative variance of the reach- and cohort-specific survival estimates. How were the variances of those reach and cohort survival estimates computed? How was the standard error on the weighted average survival computed? Provide details.*

Response: In the sub-cohort approach to estimating reach survival rates, the fish detected at LGR were stratified into a user defined number of strata (a sub-cohort is simply a stratum). The standard CJS model is used separately with those fish re-released at LGR in each stratum. Once the CJS estimates of survival are obtained, the standard theoretical variances of the CJS model, in the form of inverse relative variances as shown in the previous response are used to weight each stratum's survival rate for a particular reach, and summed to create the weighted average reach survival rates for that particular reach. The reviewer should note that the sub-cohort approach was not used in the 10-yr report. All estimates of reach survival rates are based on the CJS model applied to the full sample of fish released, rather than simply on those detected in temporal intervals at LGR.

- *In the final partial paragraph, the CSS discusses using a "per-mile" expansion of juvenile survival in cases where it was impossible to estimated survival to BON directly. Previously (Chapter 3), they used a per-km method of extrapolation. Either there or here, did they consideration other basis for extrapolation? Did they consider the goodness-of-fit of the extrapolation method used? Did they estimate the standard error on the survival estimate to BON, either with or without the extrapolation?*

Response: In the bootstrap computer program that computes all parameter estimates along with the confidence intervals, both a "per-mile" and "per-project" extrapolation is computed. The reference to "per-km" extrapolation in Chapter 3 will be revised to "per-mile" extrapolation. The rationale for choosing the "per-mile" extrapolation approach as the standard instead of the "per-project" approach has been detailed in a prior response in Chapter 3 to the same BPA comment. Goodness-of-fit was not computed. Bootstrap standard errors and confidence intervals are computed in the bootstrap computer program, and will be added to the appropriate Appendix D tables as stated in a prior response to a BPA comment on Appendix D. In the cases were an extrapolation was necessary to in order to obtain an estimate to BON, the concept of estimating a standard error on the survival rate without the extrapolation as suggested by the BPA reviewer does not make sense.

Page B-5 – BPA comment:

- *The CSS lists the three ways in which fish can pass an individual transport dam, and indicates that these three passage routes describe the passage routes through the hydrosystem. However, their three passage routes must be combined over multiple dams to describe the possible passage routes through the entire hydrosystem. For example, there are seven possible passage routes through LGR, LGS, and LMN that result in transportation from one of those dams – (i) transport at LGR (route 1), (ii) transport at LGS following either detection at LGR, or non-detection at LGS (routes 2 and 3), and (iii) transportation at LMN following either detection at both LGR and LGS, detection at only one of LGR and LGS, or non-detection at both LGR and LGS (routes 4-7). Thus, the CSS "partition" of PIT-tagged smolts arriving at LGR is, at best, unclear form their description and, at worst, potentially omitting considerable numbers of fish. Clarification in this report is required.*

Response: The CSS does not attempt to analyze all "possible" routes of passage in the manner inferred by the BPA reviewer. Instead, the CSS has created the three groupings of "possible" passage routes that best reflect what is being experienced by the untagged run-at-large. For the migration years covered in this CSS 10-yr report, the untagged run-at-large was most often transported at the three Snake River collector dams if collected there (exception is 1997 when management operations bypassed many tagged and untagged fish at LGS and LMN during parts of the migration season). We say that the collected fish were most often transported rather than 100% transported, since there are occasions over the years when all fish from raceways were returned-to-river due to unavailability of enough barges at peak passage times, or malfunctions at the facility that required short-term bypassing of all fish. Given the project operations from 1994 to 2004, the untagged run-at-large was either (i) collected and transported from one of the three Snake River transport site, (ii) collected and bypassed from one or more of these sites, or (iii) uncollected at these three sites, passing through either spill or turbines. For transported fish, the CSS utilizes those either transported from LGR, or first-time detected fish that are transported at LGS or LMN. We rely on first-time detected PIT-tagged fish at the two downstream dams, since those PIT-tagged fish match closest to the untagged run-at-large. Since we must return fish from the collector dams each year in order to estimate the inriver reach survival rates, there are occasions when these fish will be collected at the downstream sites and transported. Generally, all fish subsampled and handled in the Smolt Monitoring Program at these dams will go to transportation after handling and recovery. However, most multi-site detected PIT-tagged fish that get transported do not reflect the untagged run-at-large. Therefore, the BPA reviewer's contention that the CSS is "potentially omitting large numbers of fish" is incorrect.

Page B-6 – BPA Comment refers to "#5 Observed transportation estimate of run-at-large smolts at LGR is t_2 = (LGR run-at-large transported/LGR run-at-large collected) m_{12} and expectation of

$E(t_2) = E(m_{12})\ P_{t2}$ whereP_{t2} is the proportion of run-at-large (total fish at level of species and rearing type from Smolt Monitoring Program) transported at LGR".

- *#5. Is "run-at-large" equal to "untagged" here, or does it also include tagged fish.*

Response: The numbers of run-at-large fish collected and transported at LGR include both untagged and tagged fish. The Smolt Monitoring Program provides separate estimates of collected and transported "unclipped, non-CWT" yearling Chinook, which we use for run-at-large wild Chinook, "clipped or unclipped with CWT" yearling Chinook, which we use for run-at-large hatchery Chinook estimates, "unclipped" steelhead, which we use for wild steelhead, and "clipped" steelhead, which we use for hatchery steelhead.

- *How is P_{t2} estimated?*

Response: This parameter is an estimate of the proportion of PIT-tagged fish that would have been transported at LGR if PIT-tagged fish had been transported at the same rate as the run-at-large (see prior response for definition of run-at-large fish). It is estimated as (est. run-at-large transported)/(est. run-at-large collected) for the group of fish of interest.

- *Is P_{t2} really the proportion of the entire run-at-large that were transported at LGR, or only the proportion of the run-at-large collected at LGR that were transported?*

Response: P_{t2} = (est. run-at-large transported)/(est. run-at-large collected); therefore, it is the proportion of the run-at-large collected at LGR that were transported. We multiply P_{t2} with m_{12} to get t_2.

- *Similar comments pertain to #7 and #9.*

Answer: The same response for LGR (#5) applies to LGS (#7) and LMN (#9).

Page B-7 – BPA Comments:

- *#13 - #15: It is essential for the CSS to actually write out the expected values of the statistics T_0, T_0*, and C_1 in terms of the underlying model (i.e., survival, detection, transportation, and removal parameters), rather than leaving them partially defined. This level of technical detail is essential for all readers to know exactly what is being estimated by the parameters in the report.*

Response: The details requested by the BPA reviewer already exist in #1 to #12. In order to simplify the long formulas for the expectations of T_0, T_0*, and C_1, we feel our presentation is actually easier for readers to visualize what is being estimated. See Appendix C for formulas of expectation for T_0.

- *#15, #16: The statistics d_0 and d_1 are never defined. The 50% survival probability is not explained – 50% survival to where? On what basis is 50% chosen? Why not use the actual estimated survival probability to whatever site or sites are used?*

Response: The parameters d_0 and d_1 are defined directly below the formula and their rationale detailed in the first full paragraph on page B-8. These parameters account for PIT-tagged fish from categories C_0 and C_1, respectively, removed below LMN. Since most of this type of removal occurred at MCN in 1994 and at JDA or BON in other years, and survival from LGR to these sites was approximately 50% in the years affected, we developed into the bootstrap program a fixed 50% removal adjustment for all years. Although a year-specific estimated removal rate could have been programmed, we opted for this simpler approach when programming for this adjustment since the numbers of PIT-tagged fish affected was relatively low (numbers are presented in response to the next BPA comment).

This same basic question was raised by BPA in their review of the 2006 CSS Annual Report, and our response to them then (Berggren et al. 2006, pages 165-166) is still pertinent. "PIT-tagged fish not confirmed as being returned-to-river at a downstream dam needed to be removed from either the C_0 or C_1 study groups. Fish were considered as removals at McNary Dam when detected on the raceway or sample room monitors or only on the separator monitor during the summer transportation season, or when collected and removed at John Day or Bonneville Dam for other research purposes. Samples of CSS PIT-tag hatchery Chinook from Rapid River, McCall, and Dworshak hatcheries were collected and sacrificed at John Day and/or Bonneville dams during migration years 1999 to 2003 for physiological (blood chemistry) evaluation (Dr. Congleton, University of Idaho Fish and Wildlife Unit). Because most removals occurred at John Day and Bonneville dams for other research purposes, we settled on a fixed 50% Lower Granite to Bonneville Dam survival rate for each removed fish in order to subtract these fish in LGR-equivalents from the estimated number of smolts in Categories C_0 and C_1. Most survival rates from Lower Granite Dam to Bonneville Dam from 1995 to 2004 (excluding 2001 when extremely low in-river reach survival rates occurred on in-river migrants) have been averaging around 50%. In 1994, the wild Chinook in-river survival rate from Lower Granite Dam to McNary Dam was estimated at 47%, with virtually all removals occurring at McNary Dam since no operational return-to-river diversion system was present that year, so the fixed 50% expansion to LGR-equivalents on removals was proper in that year also. In post-1994 years, wild Chinook and wild steelhead had relatively small "raw" numbers of PIT-tag fish removed at downstream dams."

Wild Chinook	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**
Pre-adj. C0 est. #[1]	3,621	2,725	1,919	682	3,081	4,469	6,573	233	6,410	9,001
Removal # Percent	910 25.1	8 0.29	1 0.05	1 0.15	0	0	41 0.62	1 0.43	60 0.94	60 0.67
Wild Steelhead				**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**
Pre-adj. C0 est. #				454	776	1,113	1,871	103	4,107	3,343
Removal # Percent				0	13 1.68	0	0	0	9 0.22	12 0.36

Pre-adj. C0 est. # is the estimate prior to subtracting twice the removal number.

Page B-8 – BPA Comments:

- *Finally, an attempt is made at an explanation for the 50% survival probability used to deal with downstream removals. It is not sufficient, however. Why not use a dam-specific adjustment, rather than poking all downstream removals and assuming a common survival to every downstream dam? Have the effects of violations of this 50% survival assumption been examined? It is known that violations of this assumption occur, because survival between MCN, JDA, and BON is not 100%, so survival to one dam (e.g., at 50%) is not equivalent to survival to the other dams, as is implicitly assumed by using a single survival probability to all dams. Additionally, if using a single survival rates is warranted and if survival to BON is to be used each year, it should be possible to use the estimated survival to BON for the year, rather than assuming 50% survival each year.*

Response: Although dam-specific adjustments could have been used, the relatively low numbers of fish being affected as will be shown in response to the next BPA comment, makes all the concerns being raised here an over-reaction to a negligible effect.

- *Show the number of removal on a dam-specific basis that you contribute to d_0 and d_1.*

Response: The following two tables show the initial number of PIT-tagged smolts estimated in study categories C_0 and C_1 and final values obtained after the adjustment for fish removed at dams below LMN. In Table 1, the percent change from initial to final smolt estimate after the adjustment was minimal for wild Chinook, wild steelhead, and hatchery steelhead at less than 2%, except for wild steelhead in 1998 and wild Chinook in 1994. The high rate for wild Chinook in 1994 was due to no return-to-river capability that year at McNary Dam (all but two fish were MCN removals); the estimated reach survival from LGR to MCN was estimated at 47%, in line with the fixed 50% rate.

In Table 2, a higher percent change from initial to final smolt estimate after adjustment is seen for hatchery Chinook than was seen for wild Chinook or all steelhead. However, even these removal adjustment changes were generally less than 4%. The planned removals for physiological testing of PIT-tagged Chinook from Dworshak, Rapid River, and McCall hatcheries in the lower Columbia (mostly at Bonneville Dam) are the main reason for the higher percent change seen with these three hatcheries compared to Imnaha or Catherine Creek acclimation ponds. It should be noted that even if no survival rate expansion were applied, one would still, at a minimum, need to subtract the d_0 and d_1 fish removed below LMN in computing the final C_0 and C_1 smolt numbers. So relative to this minimum adjustment, the changes due to the CSS adjustments of $2*d_0$ and $2*d_1$ are effectively one -half the percentages shown in tables 1 and 2. The bottom line is that the CSS adjustment in years after 1994 has contained relatively small numbers of fish. Therefore, the suggestion of the reviewer that we should fine tune our adjustments to each dam where PIT-tag fish removals are taking place by using estimates of reach survival from LGR to that particular dam appears to be excessive. It would have relatively little effect on the resulting numbers of smolts estimated in C_0 and C_1 over the CSS approach.

Table 1. Change in C_0 and C_1 smolt estimates from initial to final value after adjusting for removals below LMN for wild Chinook and wild/hatchery steelhead.

Sp/RT Code[1]	Migr. year	Category C_0 smolt numbers				Category C_1 smolt numbers			
		final[2] C_0	initial C_0	remove d_0	change initial to final	final[2] C_1	initial C_1	remove d_1	change initial to final
WCH	1994	1,801	3,621	910	50.3%	4,431	8,459	2,014	47.6%
	1995	2,709	2,725	8	0.6%	14,206	14,260	27	0.4%
	1996	1,917	1,919	1	0.1%	5,209	5,213	2	0.1%
	1997	680	682	1	0.3%	1,936	1,936	0	0.0%
	1998	3,081	3,081	0	0.0%	12,276	12,296	10	0.2%
	1999	4,469	4,469	0	0.0%	26,140	26,150	5	0.0%
	2000	6,494	6,576	41	1.2%	16,833	17,051	109	1.3%
	2001	231	233	1	0.9%	20,307	20,589	141	1.4%
	2002	6,218	6,338	60	1.9%	12,687	12,911	112	1.7%
	2003	8,879	8,999	60	1.3%	12,694	12,846	76	1.2%
	2004	2,252	2,292	20	1.7%	16,504	16,698	97	1.2%
WST	1997	454	454	0	0.0%	2,984	2,990	3	0.2%
	1998	750	776	13	3.4%	5,150	5,374	112	4.2%
	1999	1,113	1,113	0	0.0%	6,992	6,992	0	0.0%
	2000	1,871	1,871	0	0.0%	10,616	10,616	0	0.0%
	2001	103	103	0	0.0%	11,892	11,932	20	0.3%
	2002	4,045	4,061	8	0.4%	8,726	8,802	38	0.9%
	2003	3,320	3,344	12	0.7%	7,132	7,160	14	0.4%
HST	1997	3,390	3,394	2	0.1%	19,095	19,113	9	0.1%
	1998	2,926	2,938	6	0.4%	17,958	17,998	20	0.2%
	1999	3,952	3,956	2	0.1%	20,975	20,983	4	0.0%
	2000	4,408	4,410	1	0.0%	18,804	18,808	2	0.0%
	2001	372	376	2	1.1%	19,132	19,226	47	0.5%
	2002	6,129	6,145	8	0.3%	14,038	14,110	36	0.5%
	2003	6,459	6,479	10	0.3%	10,118	10,144	13	0.3%

[1] Sp/RT is species and rear-type code: WCH = wild Chinook; WST = wild steelhead; and HST = hatchery steelhead.

[2] Final C_0 = initial C_0 - $2*d_0$ and final C_1 = initial C_1 - $2*d_1$.

Table 2. Change in C_0 and C_1 smolt estimates from initial to final after adjusting for removals below LMN for hatchery Chinook.

Sp/RT Code[1]	Migr. year	Category C_0 smolt numbers				Category C_1 smolt numbers			
		final[2] C_0	initial C_0	remove d_0	change initial to final	final[2] C_1	initial C_1	remove d_1	change initial to final
DWOR	1997	2,529	2,531	1	0.1%	3,613	3,613	0	0.0%
	1998	11,151	11,181	15	0.3%	13,128	13,214	43	0.7%
	1999	10,484	10,518	17	0.3%	19,083	19,207	62	0.6%
	2000	13,075	13,477	201	3.0%	5,416	5,580	82	2.9%
	2001	886	910	12	2.6%	16,872	17,480	304	3.5%
	2002	19,008	19,650	321	3.3%	14,914	15,570	328	4.2%
	2003	17,697	18,033	168	1.9%	6,715	6,985	135	3.9%
	2004	6,280	6,370	45	1.4%	14,009	14,195	93	1.3%
RAPH	1997	4,176	4,178	1	0.0%	6,843	6,845	1	0.0%
	1998	4,402	4,420	9	0.4%	13,597	13,691	47	0.7%
	1999	7,040	7,094	27	0.8%	14,456	14,602	73	1.0%
	2000	11,046	11,332	143	2.5%	5,248	5,406	79	2.9%
	2001	966	1,014	24	4.7%	15,989	16,631	321	3.9%
	2002	13,625	14,065	220	3.1%	14,854	15,436	291	3.8%
	2003	16,858	17,142	142	1.7%	7,055	7,195	70	1.9%
	2004	3,484	3,520	18	1.0%	12,776	12,928	76	1.2%
MCCA	1997	6,761	6,761	0	0.0%	9,272	9,274	1	0.0%
	1998	3,849	3,887	19	1.0%	12,816	12,886	35	0.5%
	1999	8,407	8,477	35	0.8%	11,391	11,527	68	1.2%
	2000	13,064	13,336	136	2.0%	4,485	4,565	40	1.8%
	2001	1,000	1,034	17	3.3%	15,536	16,040	252	3.1%
	2002	10,280	10,662	191	3.6%	12,315	12,787	236	3.7%
	2003	19,696	20,034	169	1.7%	8,669	8,817	74	1.7%
	2004	2,359	2,391	16	1.3%	16,297	16,489	96	1.2%
IMNA	1997	2,219	2,221	1	0.1%	3,785	3,785	0	0.0%
	1998	1,995	1,995	0	0.0%	6,335	6,335	0	0.0%
	1999	2,869	2,869	0	0.0%	5,084	5,084	0	0.0%
	2000	4,396	4,456	30	1.3%	2,254	2,286	16	1.4%
	2001	366	376	5	2.7%	6,939	7,043	52	1.5%
	2002	4,637	4,735	49	2.1%	5,135	5,253	59	2.2%
	2003	6,683	6,755	36	1.1%	2,908	2,936	14	1.0%
	2004	1,302	1,318	8	1.2%	4,456	4,502	23	1.0%
CATH	2001	379	391	6	3.1%	4,642	4,724	41	1.7%
	2002	2,445	2,499	27	2.2%	3,120	3,192	36	2.3%
	2003	3,201	3,247	23	1.4%	1,403	1,423	10	1.4%
	2004	503	513	5	1.9%	1,869	1,885	8	0.8%

[1] Hatchery Code is: DWOR = Dworshak; RAPH = Rapid River; MCCA = McCall; IMNA = Imnaha; and CATH = Catherine Creek.

[2] Final C_0 = initial C_0 - 2*d_0 and final C_1 = initial C_1 - 2*d_1.

- *"Estimation of SARs for study categories:" SAR1(T0) and SAR2(T0) have been discussed but not defined in the report. Define all measures.*

Response: $SAR_1(T_0)$ is a combination of dam-specific transport SARs in LGR-equivalents that are weighted by the proportion of run-at-large in total transportation occurring at each dam. $SAR_2(T_0)$ is the sum of returning adults from transported PIT-tagged fish divided by the sum of PIT-tagged smolts transported from each dam in LGR-equivalents. Parameter $SAR_2(T_0)$ is the primary SAR for evaluating transportation.

Page B-9 – BPA comments:

- *A "common annual routing rate to the raceways" was used -- what is this? What value was used?*

Response: A same rate of 2/3 PIT-tagged hatchery Chinook to raceways and 1/3 PIT-tagged hatchery Chinook returned to river for first-time detected fish from CSS participating hatcheries has been used at LGR, LGS, and LMN since 2000. It was accomplished by having the Separation-by-Code (SbyC) electronics at the Snake River collector dams divert 2 PIT-tagged fish to the raceways for every 3 PIT-tagged fish arriving from a particular CSS hatchery. In 2002 and 2003, the CSS coordinated with state and tribal researchers to divert ½ of their PIT-tagged wild Chinook to the raceways using SbyC. In 2004, this was increased to the same rate of 2/3 PIT-tagged wild Chinook and wild steelhead being routed to the raceways using SbyC. By utilizing a common annual routing rate for a group of PIT-tagged fish of interest, one achieves self-weighting across the three collector dams relative to their proportional contribution of each collector dam to total transportation. The benefit of achieving self-weighting is that $SAR_1(T_0)$ and $SAR_2(T_0)$ become equivalent in estimating the transportation SAR.

- *The notation used to define AC0 and AC1 is insufficient. It does not preclude using adults that were removed at downstream dams for any reason. Because many removed fish are not sacrificed, it is conceivable that some of these "removed" fish may return as adults. Are these adults included in AC0 and AC1?*

Response: The BPA review is mistaken. PIT-tagged smolts that are removed at downstream dams are considered permanently removed, regardless of whether sacrificed or not. For example, a fish detected only on the separator at McNary Dam later in the summer after the start of the transportation program of summer migrants would be considered as removed at that site, and therefore, any adult return from that particular fish would not be counted.

- *One assumes not, because this would positively bias the SARs for the C0 and C1 groups; however, the notation used implies that these removed fish are included in AC0 and AC1.*

Response: In the draft report we say "AC_0 = tally of adults of smolts that passed the three Snake River collector dams undetected (capture histories "1000AAAA" where A=0 signifies not being detected and A=1 signifies detection and return-to-river at a downstream site." If these fish had been removed at MCN, JDA, or BON, it would have been coded with a digit >1 in the site position of the capture-history table's field called CAPTURE_DI). Such a returning adult would not have been tallied in AC_0. This same logic applies to "AC_1 = tally of adults of smolts that passed the three Snake River collector dams with at least one detection (capture histories "11AAAAAA" or "101AAAAA" or "1001AAAA" where the A=0 signifies not being detected and A=1 signifies detection and return-to river at a downstream site. If a returning adult has a CAPTURE_DI site-position digit where A>1 in the above capture-history list, then that adult will not be tallied in AC_1.

- *It looks like SAR2(T0) is used in this report for overall SAR of transported fish, rather than SAR1(T0), unless otherwise specified. Is this correct? Clarify.*

Response: Yes, $SAR_2(T_0)$ is the primary transportation SAR parameter. Table B-1 provides a summary of which annual reports utilized $SAR_1(T_0)$ (Annual Report 2001 for wild Chinook and 2002 for both wild and hatchery Chinook) and $SAR_2(T_0)$ (Annual Report 2000, 2001 for hatchery Chinook, 2003/04, 2005, and 2006) as the primary measure of transportation SAR. The clarification of why we returned in 2003 to using $SAR_2(T_0)$ as the primary transport SAR is detailed form the bottom of page B-8 through top of page B-9 in Appendix B of the 10-yr report.

Page B-11 – BPA comments:

- *The expected value of the size of the C1* group should be presented. At the least, the definition of the C1* group should be explained. It does not make intuitive sense to define it in terms of the T0, C1, and T0* statistics, because the T0 and T0* statistics are based on different groups of fish.*

Response: Contrary to what the reviewer suggests, the parameter T_0 and T_0* are based on the same underlying group of PIT-tagged fish. When this group if PIT-tagged fish are expanded to LGR equivalents and summed, we get the starting number of smolts in group T_0 at LGR. Further, expansion of this group allows us to estimate the number that would have been in T_0, which we call T_0*, provided T_0 fish had been transported at the same rate as the untagged run-at-large. In that situation, the population arriving LGR forebay of PIT-tagged fish of a particular CSS group, such as Rapid River Hatchery Chinook for example, would consist of C_0 fish "destined" to pass three collector dams undetected, T_0* fish "destined" to be collected and transported, and a remainder of fish that are "destined" to be collected and bypassed assigned to group C_1*. The sum of the T_0 and C_1 fish equal the collected portion of the PIT-tagged group. By subtracting the number of fish in T_0* from the sum of T_0 and C_1, we obtain an estimate of residual bypassed fish. In most years this is a very small, often immeasurable number, but in 1997 when the management action was to route many untagged fish, this group accounted for upwards of 25% of the run-at-large population of Chinook and steelhead.

B-12 – BPA Comments

- *The CSS states that "the rate of harvest is assumed independent of whether fish had been transported as smolts. [These] assumptions ... apply to both TIR and D." Where does the CSS actually make use of this assumption? Is it only in their interpretation of results about TIR and D?*

Response: This assumption about harvest rate is utilized primarily when addressing losses during the adult fishes' upstream migration from Bonneville Dam to Lower Granite Dam. Although the rate of harvest is likely unaffected by whether smolts outmigrated in barges or inriver, the opportunity for harvest as transported fish may experience more straying effects could be another reason why we observed differential loss during the upstream migration based on prior downstream migration history. But even in the lower Columbia River prior to passing Bonneville Dam, there are opportunities for harvest in some years which we cannot directly measure with the PIT-tag data. Here again, we assume the rate of harvest is independent of prior downstream migration history. The effects of harvest removal will be to lower the magnitude of estimated SARs of both inriver and transported fish, but it will have less of an effect on those parameters that are based on the ratios of these two SARs (e.g., *TIR*s and *D*) if the harvest rates are independent of downstream migration history.

Reference:

Zar, J.H. 1984. Biostatistical analysis, 2nd Edition. Prentice-Hall, Englewood Cliffs NJ. 718 pp.

Department of Energy

Bonneville Power Administration
P.O. Box 3621
Portland, Oregon 97208-3621

Environment, Fish and Wildlife

July 31, 2007

In reply refer to: KEW-4

Michele DeHart
Fish Passage Center
1827 NE 44th Ave., Suite 240
Portland, Oregon 97213

Dear Ms. DeHart:

Thank you for the opportunity to comment on the Comparative Survival Study (CSS), Ten-year Retrospective Analyses Report, May 30, 2007. We have included points below and provide a more detailed analysis as enclosures including a (1) General Technical Comment; (2) Evaluation of the CSS Response to the ISAB Recommendations; and (3) Detailed technical review comments.

The (CSS) 10-year Retrospective Report provides a history of PIT-tagged salmonid fish performance from 1998 through 2006. The length of time, the breadth of geographic coverage, and range of salmon life-history phases investigated in the report have the potential of providing a valuable chronicle of recent Columbia River trends. No other study in the Fish and Wildlife Program has the same scope of effort.

As we have emphasized in past reviews of CSS Annual Reports, and now for this CSS Ten-year Report, an overriding issue for CSS analyses is reproducibility. It is imperative that CSS analyses be capable of accurate reproduction or replication by independent researchers to see if their analyses give similar results to those reported by the original group. The ability to reproduce results is crucial to the scientific review process. Reproducibility requires transparency in terms of sufficient data and detailed methods to allow a third party to reproduce the analyses contained in the Report. As has been noted in the past (e.g., the ISAB 2005 CSS review, in the review by the ISAB on the 2007-2009 CSS Proposal, and BPA's Review of the 2005 Annual Report), CSS analyses have not always been sufficiently transparent. The CSS Ten-year Retrospective Analyses Report continues this pattern, as it does not include sufficient data and detailed methods to allow a third party to reproduce the analyses and conclusions contained in the Report.

- Tagging Results and Reproducibility -- Our attempts to reconstruct final results from intermediate calculations presented in the report have been limited by the absence of necessary information or insufficient technical description.

- Latent Mortality -- When the performance measure for "differential mortality" is corrected for the extra migration of upstream stocks, there is little or no evidence of differential hydrosystem mortality for hatchery Chinook salmon.

- Tagging Results --The CSS Report needs to simply document and display the tagging results for the benefit of most readers and organizations that do not have the resources of the CSS organization. This issue is fundamental to our comment - the need to provide the means to reproduce results.

- Non-standard practices -- The report includes non-standard modeling practices resulting in limited use of the analyses. These practices need to be peer reviewed.

- Missing information -- Basic information and mathematical definitions for equation parameters such as SARs, and also the number of fish actually transported at each dam are absent.

- Upstream and downstream comparisons -- CSS continues to compare upstream and downstream Chinook salmon stocks when the data clearly do not support such comparisons. Previous critique of the upriver-downriver comparison including the 2005 ISAB review has documented this point. The CSS Report does not demonstrate a biological difference given fish size, migration date, marine arrival timing, and year, in addition to upstream/downstream classification.

- Invalid assumptions -- The analyses assumes that no natural mortality occurs once salmon pass the first upstream dam, thus concluding that all mortality between upstream and downstream dams is caused by the hydrosystem. When the performance measure is corrected for the extra migration of the upstream stocks, there is little or no evidence of differential hydrosystem mortality for hatchery Chinook salmon.

- Due to the inability to reproduce these results using accepted modeling and analytical procedures the CSS Report's findings do not demonstrate the scientific rigor and support to authoritatively guide hydrosystem management.

Please let us know if you have any questions or require further clarification on our comments. As we stated in our 2005 and 2006 comments on the CSS Annual Reports, it is critical that the issues raised be addressed because of their importance for the continuing work under the project.

Sincerely,

/S/ Robert J. Austin

Robert J. Austin
Deputy Director of Fish and Wildlife

Enclosures

cc:
Dr. Tom Karier, Northwest Power & Conservation Council
Mr. Bill Booth, Northwest Power & Conservation Council
Mr. Jim Kempton, Northwest Power & Conservation Council
Ms. Joan Dukes, Northwest Power & Conservation Council
Mr. Bruce Measure, Northwest Power & Conservation Council
Ms. Rhonda Whiting, Northwest Power & Conservation Council
Mr. Larry Cassidy, Northwest Power & Conservation Council
Ms. Melinda Eden, Northwest Power & Conservation Council
Mr. Tony Grover, Northwest Power & Conservation Council
Mr. Brian Lipscomb, Columbia Basin Fish & Wildlife Authority
Chairperson Wanda Johnson, Burns Paiute Tribe
Chairman Chief Allen, Coeur d'Alene Tribe
Chairman Michael Marchand, Confederated Tribes of the Colville Reservation
Chairman Glen Nenema, Kalispel Tribe
Chairperson Jennifer Porter, Kootenai Tribe of Idaho
Chairperson Samuel Penney, Nez Perce Tribe
Chairman James Steele Jr., Confederated Salish & Kootenai Tribes
Chairperson Alonzo Coby, Shoshone Bannock Tribes of Fort Hall
Chairman Kyle Prior, Shoshone Paiute Tribes of the Duck Valley Reservation
Chairman Richard Sherwood, Spokane Tribe of Indians
Chairman Antone Minthorn, Confederated Tribes of the Umatilla Indian Reservation
Chairman Ron Suppah, Confederated Tribes of the Warm Springs Reservation
Chairman Lavina Washines, Yakama Indian Nation
Ms. Mary Verner, Upper Columbia United Tribes
Mr. Olney Patt Jr., Columbia River Intertribal Fish Commission

bcc:
S. Wright - A-7
L. Bodi - A-Seattle
S. McNary - A-7
K. Hunt - DKR-7
B. Cobell - DKT-4
K. Johnston - DKT-7
R. Bennett - K-7
S. Stewart - KEC-4
B. Austin - KEW-4
B. Maslen - KEW-4
G. Dondlinger - KEWB-4
P. Lofy - KEWL-4
K. Fisher - KEWR-4
P. Krueger - KEWR-4
A. L'Heureux - KEWR-4
K. Powers - KEWR-4
M. Shaw - KEWU-4
L. Grimm - LC-7
P. Key - LC-7
Official File - KEW (FW-24-11)

Austin:mpr:4748:07/24/2007 (KEW-4-W:\Kew\KEW07\FW-24-11\CSS Comment Letter.doc)

Review of

COMPARATIVE SURVIVAL STUDY (CSS) of PIT-Tagged Spring/Summer Chinook and Steelhead in the Columbia River Basin: Ten-Year Retrospective Analyses Report

by

Bonneville Power Administration
P.O. Box 3621
Portland, Oregon 97208-3621

July 24, 2007

Table of Contents

General Comments

The Comparative Survival Study (CSS) 10-year retrospective analysis provides a useful history of PIT-tagged fish performance from 1998 through 2006. The length and breadth of the tagging data provide a valuable look at the history of salmonid survivals, travel times, transport/inriver ratios (TIRs), smolt adult returns (SARs), etc., in the Columbia Basin. No other study has the same temporal, geographic, or salmonid life-history scope as this project. For this reason, documenting the data collected and the status and trends of the estimate of various performance measures is crucial for the Columbia Basin Fish and Wildlife Program which has supported this work. It is therefore surprising that this important task is limited to a relatively few tables and graphs in Chapters 2 and 3. Appendix D supplements the information in these chapters but never quite reaches the level of showcasing the important trends in the results. In many cases, standard errors or confidence intervals are neither reported nor displayed.

As urged by the Independent Scientific Advisory Board (ISAB), this CSS report now presents some of the methods used for estimating SAR, TIR, and latent mortality (D) for the various groups of interest (i.e., the C0, C1, and T0 groups). It has compiled methods strewn throughout previous reports in one place, and this makes reading the report much easier. Nevertheless, this encouraging start was not continued throughout the report or consistent across chapters. Complicated performance measures such as annual SARs are described verbally but never mathematically defined in equations. Cryptic tables are included, showing values used to estimate annual SARs (Tables D-20 and D-30), but it is not clear what these values are or how they were combined to estimate annual SAR. In a report as important as this 10-year review, the first priority should have been simply presenting the facts (i.e., results). Closely tied to this first objective should have been much more transparency and clarity of methods in this report. Attempts to reconstruct final results from intermediate calculations have usually been difficult due to lack of necessary information or insufficient guidance. The ability to reproduce results is crucial to the scientific peer review process.

A large proportion of the 10-year review focuses on interpreting the PIT-tag results and assessing the influence of environmental and hydrosystem effects on inriver survival and adult returns. The 10-year review includes both approaches previously described in annual reports as well as new analytical methods. These analyses are both the most interesting and often problematic aspects of the report from an analytical perspective.

In Chapter 2, the concept of instantaneous mortality rate (Z) is introduced. However, it is not based on failure times (i.e., death times of PIT-tagged fish) as it properly should be but, rather, on a simple function of the ratio of reach survival estimates and median travel times. The authors then go on to analyze survival, travel times, and Z as if they are three independent pieces of information. Reach survivals within a season are relatively stable while travel times show marked seasonal trends. Using the ratios of this information, Z values are calculated and seasonal trends are (mis)interpreted as survival processes. In fact, the trends in Z are nothing more than inverse trends in travel times misinterpreted or misconstrued as survival effects. Curiously, results of model analyses on reach survivals are not discussed, leaving the impression that results of Z values are applicable to S, which is *not* true. Finally, the authors interpreted the instantaneous mortality rate (Z) as the probability of mortality (i.e., $1-S_t = 1-e^{-Zt}$).

As requested by the ISAB, the CSS has compiled in Chapter 3 and Appendix B many of the methods used to generate the time series of estimates reported. Nevertheless, some methods and definitions are missing here and throughout the report (e.g., annual SAR). Also missing are certain basic results, such as the number of fish actually transported at each dam, which should be documented in this report. It is very helpful to see the figures of trends in the performance measures over time, and to see comparisons between hatchery and wild stocks. Also the 90% confidence intervals included on some figures aid interpretation. However, the CSS Report has based too much inference on whether confidence intervals on two estimates overlap. Non-overlapping confidence intervals is an invalid test of significant differences. Instead, the CSS should find valid methods of testing significance, either within their bootstrap approach or separately with a parametric approach.

Chapter 4 explores the causes and nature of the interannual variation in performance measures such as SAR, TIR, and D. Multiple regression was used to model the responses. Although the summary tables are cryptic, it appears models with partial regression coefficients had signs inconsistent with the investigation philosophy (e.g., negative sign with flow) were consistently omitted. This practice left models that had interaction terms but no main effects or quadratic terms without the linear component inconsistent with general model building practices. Another aspect of the report used parametric models to partition total variance of metrics into natural variation and measurement error. However, the assumption, for example, that SARs are binomially distributed is inconsistent with the mark-recapture models used to estimate the values. Underestimating sampling error will positively bias estimates of natural variation. The report needs to use goodness-of-fit tests to assess the model assumptions and compare their parameter estimates with those of the nonparametric variance component formulas provided. Their inferences concerning natural variation do not take into account their own findings on ambient effects, the historical distribution of those factors, or how influences such as global climate change many affect projections in the future.

The CSS continues in Chapter 5 its comparison of upstream and downstream Chinook salmon stocks. As in the past, multiple upstream hatcheries and collection points are used, while only a single downstream hatchery and collection point (for wild fish) is used, despite the ISAB's recommendation to incorporate more downstream stocks. Given that this is a retrospective report, it is understandable that the CSS could not immediately include additional downstream stocks. While the CSS does perform useful comparisons of biological characteristics of the upstream and downstream stocks, their upstream-downstream analysis is invalid in other critical ways. The CSS uses an invalid performance measure to identify delayed mortality caused by the hydrosystem. This approach assumes no natural mortality for smolt should occur between upstream and downstream sites. When the performance measure is corrected for the longer migration of the upstream stocks, there is little or no evidence of delayed hydrosystem mortality for hatchery Chinook salmon. Similarly, the CSS Report does not consider the longer distance to travel for upstream stocks when comparing travel and arrival times of upstream and downstream stocks. Even if the hydrosystem were not in place, the upstream stocks would still have farther to travel than downstream stocks.

Chapter 6 attempts to partition survival across different portions of the migration, focusing on smolt survival from the hatchery/trap to LGR, perceived adult survival from BON to LGR, and perceived adult survival from LGR back to the hatchery/spawning grounds. Adults are categorized by juvenile migration method. The effect of juvenile migration method—in particular, transportation—on perceived adult upriver survival is an important question, and the analyses in this chapter relating adult survival to migration method are worthwhile. The CSS should provide the methods used in estimating upriver survival for a given juvenile release group. Reviewing and reproducing their results is difficult without those methods. Additionally, the report misinterprets the odds ratio from their logistic regression when comparing adult survival for LGR-transport fish to other fish; consequently, they overestimate the effect of LGR transportation on upriver adult survival.

Chapter 7 describes simulations done to assess the effect of violations of key Cormack-Jolly-Seber (CJS) assumptions on estimation of C0, T0, C1, SARs, TIR, and D. The assumptions considered were (1) all fish have common survival and detection probabilities, and (2) detection has no effect on subsequent survival. Assumption violations considered were (1) temporal changes in survival and detection probabilities, and (2) differential inriver survival of pre-assigned groups (T=transport group, R=return-to-river group) based on past detections. This is an important exercise, demonstrating the robustness of the estimation methods to all but severe temporal changes in survival and detection probabilities, and the dependence of estimation methods on the assumption of common survival regardless of past detections. The focus on the T and R groups is reasonable, given the ISAB recommendation to pre-assign future transport groups in this way. However, the assessment of assumption violations using the T and R groups does not translate directly to the C0, C1, and T0 groups or to the study design used in the past. The CSS should have performed a third set of simulations assessing the effect of detection-influenced survival directly on estimates of C0, C1, and T0, in order to more correctly assess the robustness of past analyses.

Chapter 8 provides a summary of objectives and findings from the 10-year retrospective report. Because conclusions are at times based on the invalid analysis of the earlier chapters, their inferences are invalid as well. The CSS attributes all differences in survival and travel time between study groups in the upstream-downstream comparison to the hydrosystem, ignoring differences expected because of different migration distances. Additionally, the CSS incorrectly claims to have addressed the question of whether smolt transportation compensates for effects of the FCRPS on survival of Snake River Chinook and steelhead. At most, the comparison of the SARs of transport fish and inriver fish indicates whether transportation is a viable management option; it is not equivalent to comparing transportation to migration through the unimpounded river. The question of the effect of the FCRPS on salmonid migration and survival is important. However, it is not addressed by the analyses presented in this report.

Review of Chapter 2

In this chapter, travel time, survival, and a measure of instantaneous mortality were estimated over two reaches, LGR–MCN and MCN–BON for the years 1998-2006 for hatchery/wild yearling Chinook salmon and steelhead. Multiple regression analysis was used to examine the relationship between these metrics and various environmental covariates. Within season, eight weekly cohorts were formed to monitor trends within the year.

- Page 18, last paragraph

The report used the exponential decay model

$$N_t = N_0 e^{-Z_t} \quad (1)$$

to derive a measure of instantaneous mortality rate Z. Solving for Z in Eq. (1) yields

$$Z = \frac{-\ln\left(\frac{N_t}{N_0}\right)}{t}$$

or

$$Z = \frac{-\ln\left(S_t\right)}{t}. \quad (2)$$

The report then goes on to estimate Z by

$$\hat{Z} = \frac{-\ln\left(\hat{S}\right)}{\hat{F}TT} \quad (3)$$

where

$\hat{S}$ = reach survival rate,

$\hat{F}TT$ = median fish travel time for the fish that survived the reach.

Equation (3) provides a convenient but biased estimate of the instantaneous morality rate. Properly, the maximum likelihood estimate of Z would be based on the likelihood model

$$L = \prod_{i=1}^{N} Z e^{-Zt_i}$$

and estimator

$$\hat{Z} = \frac{1}{\bar{t}} \quad (4)$$

where t_i = lifetime for the ith fish $\left(i = 1,\ldots,N\right)$. However, PIT-tag data do not provide lifetimes for the fish, only travel times for the survivors. Therefore, the PIT-tag data are incapable of estimating instantaneous mortality rates. Any relationship between the true estimates of Z [Eq. (4)] and that used in the report [Eq. (3)] may be appropriate at best and seriously biased as worst.

- Page 20, second paragraph

In performing the regression analyses, the response variables were

a. $\ln\left(\hat{S}\right)$

b. Median ln (FTT)

c. $$Z = \frac{-\ln(\hat{S})}{FTT}$$

or

d. $$\ln \hat{Z} = \ln(-\ln \hat{S}) - \ln FTT$$

Both weighted and unweighted regressions were performed using a variety of weights:

a. Inverse variance

b. Inverse CV

c. Inverse CV^2

The report seemingly takes a shotgun approach to the analysis. In the results section, which weighing scheme and why its selection was not revealed. The weight selection should be objective. Proper weighting should be inversely proportional to the variance except when the variance estimates is correlated with the response variable. In this case, the weight should be inversely proportional to the variance but adjusted to eliminate the correlation.

In the case of $\ln(\hat{S})$

$$\mathrm{Var}(\ln \hat{S})\ \mathrm{B}\ \frac{\mathrm{Var}(\hat{S})}{S^2}.$$

However, $\mathrm{Var}(S)$ in a CJS model is proportional to S, saying $\mathrm{Var}(S) = S \cdot f(n)$ where $f(n)$ is a function of sample size and detection probabilities. Then

$$\mathrm{Var}(\ln \hat{S}) = \frac{Sf(n)}{S^2} = \frac{f(n)}{S}.$$

Consequently, the proper weight should be inversely proportional to that quantity after adjustment for S, where

$$W = \frac{1}{\left(\frac{f(n)}{S}\right)} \cdot \frac{1}{S} = \frac{1}{f(n)},$$

or in other words,

$$W = \frac{\hat{S}}{\widehat{\mathrm{Var}}(\hat{S})},$$

which was *not* one of options considered by the CSS report.

As Z is estimated in the report,

$$\hat{Z} = \frac{-\ln(\hat{S})}{FTT},$$

analyses of $\hat{S}$, FTT, and Z are not independent. For example, by the formulation of Z, if the FTT have a downward seasonal trend and $\hat{S}$ is static, then Z will have an upward seasonal trend (e.g., Fig. 2.4). There is no new information conveyed by the third relationship that is not known for the first two trends. Only if Z was actually estimated by actual fish lifetimes would it provide new information not already captured by $\hat{S}$ and FTT.

- Page 20, last paragraph

The report states, "we examined the sign of the parameter coefficients for plausibility and eliminated models with implausible sign." This is a dangerous and potentially misguided approach to modeling. First, such an approach eliminates the possibility that new insights might be developed and assumes all preconceptions are correct. Secondly, it is unwise to directly interpret the sign (+ or −) of partial regression coefficients (Neter et al. 1996:290-291). Such signs do not necessarily indicate a positive or negative relationship between dependent and independent variables but, instead, adjustments of the model in the presence of other covariates. This unorthodox model strategy can lead to odd modeling results (see comments below).

- Page 20, last paragraph

The report states, "models were fit and ranked according to their AIC_C and BIC scores." However, many tables (e.g., Tables 2.7-2.11, 2.13) report AIC scores while other tables (e.g., Tables 2.12, 2.15-2.16) report AIC and AIC_C scores. What was actually done and reported needs to be clarified. For example, are the AIC values in Tables 2.7-2.11 actually AIC_C and "AIC" is a typo?

- Page 21, Section "Comparing survival modeling approaches," first paragraph

"Integrated models of fish travel time and instantaneous mortality, with each component modeled being a function of environmental covariates" are mentioned but never described. If a multivariate computational model was actually used, it needs to be provided, along with associated assumptions (providing Eq. 2.2 is inadequate).

- Page 22, multiple references on this page

Julian day was found in several instances to help describe regression relationships. The implication of this covariate in the models must be described for it is unlike the other covariates considered (e.g., WTT, percent spill, etc.). Julian date is a surrogate for numerous factors that may have a within-season trend including smoltification, flows, temperature, turbidity, etc. If the purpose of the regression analyses is to describe environmental and hydrosystem factors affecting fish response, inclusion of Julian data obscures the results. In some instances, (e.g., Table 2.15-2.16), it does a very good job all by itself!

- Page 23, fifth paragraph – Comparison of survival modeling approaches

AIC scores cannot be compared across different data sets (Burnham and Anderson 2002:80-81). Comparison of models of FTT and instantaneous mortality versus direct survival is inappropriate and Table 2.2 should be eliminated from the report.

- Page 24, first paragraph

The authors are totally misinterpreting their estimates of instantaneous mortality Z. In this paragraph, they are equating Z to probability of mortality which is *wrong*. For example, the instantaneous rate of 0.112 (steelhead, MCN–BON) is equivalent to a daily survival probability of

$$S = e^{-0.112} = 0.8940$$

or mortality of 0.106, not 0.112 as reported. A half-day has a survival probability of

$$S = e^{-0.112(0.5)} = 0.9455$$

or a mortality probability of 0.0545, not 0.056 as reported. The rest of the paragraph has similar problems and needs to be corrected. The reported values are vaguely close to the actual values only because

$$1-e^{-x} \mathrm{B}\, x \text{ for } x \leq 0.10$$

in a Taylor series expansion.

- Figs. 2.1, 2.2, 2.7

  The symbolism for box and whisker plots is not universally consistent or known. Captions should explain the symbolism.

- Table 2.1

  Caption fails to indicate which models the results refer to.

- Table 2.2

  Omit because AIC are not comparable across different datasets.

- Tables 2.7-2.16

  Captions are inadequately described. Symbols for models are cryptic and need to be explained for clarity of interpretation.
  The selection of models examined is at times eccentric:

  1. Models may include an interaction term without one or both of the main effects included. Purpose of an interaction term is to modify the main effects; it is unclear what the interaction term means in the absence of the main effects.

  2. Higher-order polynomial terms are included in models without corresponding lower-order terms, which is not conventional in linear models; for example, squared term without the linear term.

  Wonder whether this nonconventional approach to modeling is a direct consequence of dropping factors that are perceived to have the wrong sign for the partial regression coefficient (see comment above).

- Fig. 2.17

  The 20-day curve should be eliminated because the model is extrapolated beyond the range of the data. Fig. 2.1 indicates water transit time in LGR-MCN rarely if ever reaches 20 days.

Review of Chapter 3 and Appendix D

Chapter 3 and Appendix D present results on SARs, TIRs, and D for wild and hatchery spring/summer Chinook salmon and steelhead. Point estimates are presented and, in many instances, 90% bootstrap confidence intervals. The point estimates of SAR, TIR, and D are summarized by the geometric mean. Comparisons are made across migration groups, rearing types, and years. The estimated values of SAR, TIR and D are compared to benchmarks (i.e., 2% and 4% for SAR, 1 for TIR and D).

Page 51 (lines 24-26, 33-36); page 63, Table 3.2; page 74, Table 3.4 – Hydrosystem survival and system survival

- "Hydrosystem survival" includes indirect mortality effects of hydrosystem, despite the ISAB's recommendation to stop focusing on latent mortality because of the inability to estimate indirect mortality effects of the hydrosystem.

- "Hydrosystem survival" and "system survival" can be >1, and so are not actual survivals. At the very least, both performance measures are misnamed, and should not be used for management discussions.

- No benchmarks or target values for hydrosystem survival or system survival are given for comparison to estimated values. No expected values are given. Without this information, it is impossible to use the estimated values of these performance measures for management.

- "Hydrosystem survival" is introduced on page 51, defined formally on pages 59-60, and then not used because it cannot be estimated. Instead, "system survival" is reported.

- 2001 has a value of system survival of 2.139 (Table 3.2), which is >1; very high "system survival" in a very low flow year, which generally had poor inriver survival ($\hat{S}_R$ = 0.25 for 2001 [Table D-31]). It is not clear how to interpret this reported result. This result suggests the way the report is estimating system survival is invalid.

- System survival is mostly >1 for wild steelhead (Table 3.4), again inconsistent with general knowledge.

- Values of system survival are not given for hatchery Chinook salmon and steelhead.

Page 51, lines 30-42

The assumption of no density-dependent mortality needs support and should be included here. It has been hypothesized that one way in which hatchery fish negatively impact wild fish is through density-dependent mortality in estuary and nearshore ocean environments, by attracting more predators and competition for resources (food, shelter).

Pages 61-79

- 90% confidence intervals on some (but not all) performance measures were found using bootstrap methods. It is commendable that confidence intervals were computed for the performance measures, because it is impossible to interpret point estimates alone. However, it has been found (Lowther 2002) that bootstrap confidence intervals are not superior to theoretical normal theory confidence intervals arising from mark-recapture data analyzed with the Cormack-Jolly-Seber (CJS) model.

- Report all performance measures with confidence intervals, including:

  - The geometric means of the observed SARs, TIRs, or D values over the years of the study.

- Annual “system survival” estimates.

- Annual extrapolated estimates of inriver survival (S_R) from LGR to BON (Tables D-21 to D-28).

- Bootstrap confidence intervals do not easily yield confidence intervals or standard errors on performance measures that are functions of other parameters. Rather than report measures without some accompanying measure of uncertainty, standard errors or confidence intervals should have been computed in some way:

    - Geometric mean SAR, TIR, or D: A standard error for a geometric mean can be easily derived, assuming $\ln x_i$, nominally distributed, and using the expression for a geometric mean of

    $$\overline{x}_{\text{GM}} = e^{\left\{\frac{1}{n}\sum_{i=1}^{n}\ln x_i\right\}}.$$

    Otherwise, arithmetic means should be reported for they always provide an estimate of expected value.

    - Extrapolated (“expanded”) S_R , inriver survival from LGR to BON: The extrapolated S_R is a function of CJS survival estimates and river km, and a standard error for $\hat{S}_R$ could easily be found using the delta method and CJS-based variances and covariances.

    - System survival, defined in terms of inriver survivals and project-specific D: A bootstrap confidence interval could be found for system survival but would require computing system survival for each bootstrap iterate, as was apparently done for TIR and D.

  Again, standard errors or confidence levels should be computed for all performance measures and included in this report.

- Significant differences in point estimates are incorrectly based on non-overlapping 90% bootstrap confidence intervals.

    - It is possible that two estimates with overlapping confidence intervals are statistically significant.

    - One reason is that confidence intervals ignore the possible correlation between the measures being compared, e.g.,

        - SAR(T0) and SAR(C0) are correlated for a single data set, because both T0 and C0 are estimated using the same CJS parameter estimates

    - Another reason is unequal variances.

    - Even if overlapping confidence intervals were an appropriate gauge of statistical significance for SAR(T0) and SAR(C0), this method focuses on the *difference* between SAR(T0) and SAR(C0), whereas the appropriate measure is their *ratio*, or TIR.

  Therefore, determination of significant differences should be recalculated based on formal statistical tests, not on whether confidence intervals overlap.

Page 54; lines 34-41, 58-59

- S_R , inriver survival from LGR to BON, is extrapolated (“expanded”) to BON on a per-km basis in cases where it is not possible to estimate it directly using the CJS model. This is reasonable, but it should be recognized that this is extrapolation past the available data, not simply an “expansion.”

- It is unclear if other methods of extrapolation were considered, such as pre-project, and if the goodness-of-fit of the extrapolation was considered. [Should these be considered for this report?]

- Again, confidence intervals or standards errors need to be calculated and included in this report.

  - Confidence intervals are not shown on any estimate of S_R , extrapolated or not, in Figs. 3.8 (p. 68) to 3.18 (p. 77).

  - No measure of uncertainty (e.g., standard error or confidence interval) is provided for the extrapolated S_R point estimates (Tables D-21 to D-28).

Pages 58 and 63, lines 20-22

- Steelhead jacks are included in SARs, but not Chinook jacks, because

  - Steelhead jacks have a fairly stable rate of return.

  - Chinook jacks have a variable return rate.

  Removing jacks from the analysis because of their questionable contribution to spawning is understandable, but not because of a “highly variable jack return rate” (p. 63).

Page 58, lines 16-18

- It appears that conclusions (about D averaging 0.5 for hatchery and wild Chinook “in recent years”) are being presented prematurely and inaccurately in the methods section. These statements should be removed from the methods and included, with corrections and justification, in the discussions section.

- Based on CSS estimates of D for hatchery Chinook reported in Tables D-22 through D-26, only 3 of 36 point estimates for D were at 50% or less.

Page 58, lines 26-34

- The measures S_R and S_T are called “hydrosystem survival,” but these are *not* the hydrosystem survival described on pages 51, 59, 60. Please explain or clarify.

Page 59, lines 13-21

- Measures SAR_{T1}, SAR_{T2}, SAR_{T3}, SAR_{C1}, SAR_{C2}, SAR_{C3} need to be defined formally using both verbal and mathematical expressions.

- New notation is used and needs to be defined; C1, C2, C3, T1, T2, T3.
  - C1 is previously used a different context, apparently.

Pages 61-78; Appendix D

- Neither the actual numbers of tagged smolts transported from each dam nor the sample sizes used in the analyses are reported. This information needs to be included for a complete and accurate peer review.
- It is unreasonable to assess the effectiveness of transportation based on small transport groups, even if they are augmented by the LGR equivalent approach.
- Present project-specific TIRs; they are used in estimation of annual SAR in the body of the report but are not specified.

Page 61 (lines 26-28); page 68 (lines 14-15)

- 2004 returns are incomplete, so it is unreasonable to say that 2004 SAR is “low” at this point.

Page 61, lines 30-32

- The ISAB review of the 2005 CSS Annual Report referred to the NOAA finding that PIT-tagged survival < untagged survival. If the NOAA finding is true, then comparing SARs from PIT-tagged fish to target values is unreasonable unless we know the size of the bias introduced by tagging or tag loss.

Page 62 (line 13); Table D21

- The geometric mean is used to summarize point estimates of SAR, TIR, and D across years.
  - Use of the geometric mean needs justification, especially considering past criticism and the fact that the geometric mean will always be lower than the arithmetic mean.
  - Standard errors or confidence intervals need to be reported for the geometric mean (see earlier comment and suggestion).
  - Low precision on D and TIR casts doubt on conclusions based on the geometric mean, especially those based only on a point estimate.
  - The geometric mean inherently dampens the effect of extreme values, so the policy of excluding 2001 values from the geometric mean needs further justification.

Page 66, 70, D-17 (Tables D-29, D-30) – Annual SAR

- Annual SAR is discussed often and is described in words, but is never defined formally. An equation is needed to see exactly how the various components are incorporated.
- Annual SAR values should be reported in a table for all species and stocks, with confidence intervals or standard errors.

- It would be useful to compare the annual SAR values to a simple ratio of the number of adults at LGR divided by the number of juveniles at LGR.

- Tables D-29 and D-30

  - These tables should be explained clearly in the text, using precise equations and clear definitions of notation.

  - It is unclear how the values reported here were defined, estimated, and used to compute the annual SARs.

  - It is unclear what the S's mean, and what reaches they apply to.

  - It is unclear where the covariances come from.

  - No comparable tables were provided for hatchery fish.

Page 67, Figure 3.7; Page 70, last paragraph

- Figure 3.7 shows that the trend in SAR for wild fish over two- or three-year time periods mimics the trend in SAR for certain hatchery stocks. However, Fig. 3.7 also shows that SAR for wild fish did not closely track SARs for any single hatchery throughout the entire time period considered. It is therefore uncertain which single hatchery could be used as a surrogate for wild fish in future years.

- No error bars are provided on Fig/ 3.7.

Figures 3.7, 3.8, 3.9, 3.10, 3.13, 3.14, 3.15, 3.18, 3.19, 3.20 (pp. 67-78)

- Need confidence intervals or standard errors on these graphs.

Chapter 4 Review

Chapter 4 attempts to estimate environmental stochasticity in SARs, TIRs, and D by removing variability in estimates due to measurement error. Parametric methods based on beta-binomial random variables and the lognormal distribution are used. Beta and lognormal probability distributions meant to describe variability in SAR, TIR, and D due to environmental stochasticity are presented.

Page 81

- 1st paragraph – Estimates of SARs are also indicators of inriver conditions, fish health, ocean conditions, and harvest survival.
- 2nd paragraph – Opportunistic sampling of fish, more than increasing variance, may result in biased estimates.
- 3rd paragraph – As the SARs are calculated in Chapter 3, they certainly do not have a binomial sampling variance, for both the numerator and denominator (i.e., C_0 fish) are estimated random variables. For a binomial variance to be true, the denominator of the SAR would have be known without error.
- 4th paragraph – The belief that there is a single probability distribution of SAR, TIR, or D over a long time period assumes that there is no temporal trend in the measures, such as may be caused by global climate change. Chapter 3's focus on trends in these measures suggests an assumption that the measures are changing over time, which is inconsistent with the assumption that they arise from a single beta distribution.

Page 82, lines 15-17

- Akcakaya (2002) is cited as a foundation for the method used to remove sampling variance to estimate environmental variance. The method presented in Akcakaya (2002) is appropriate for census data, but not for mark-recapture data, such as the data analyzed in this report. Akcakaya (2002) refers to both Burnham et al. (1987) and Gould and Nichols (1998) for variance-components method of removing sampling variance from mark-recapture data (see below, comment on pages 82-87).
- The methods used in this chapter are not clearly presented, either in the chapter or elsewhere in the report, despite the ISAB request that the report present all methods. They are presented verbally, but not mathematically.

Page 86, lines 19-22

- Equation(4.4) for the variance of a product applies only for independent random variables. This equation cannot be used to calculate the variance of a product of inriver survivals over adjacent reaches (i.e., S_R), because these survival estimates are correlated as based on the CJS model. Instead, the delta method (Seber 1982:7-9) should be used.

Page 82-87

- Kendall's (1998) method is a parametric approach to variance component estimation that makes unnecessarily restrictive assumptions, i.e.,

  a. Measurement error is binomially distributed.

b. SARs are beta-distributed.

Extension of the method to include log-normal distributions is also unnecessarily restrictive.

- Using the conditional variance formula

$$\mathrm{Var}\left(\hat{\theta}_i\right) = \mathrm{Var}_2\left[E_1\left(\hat{\theta}_i \middle| 2\right)\right] + E_2\left[\mathrm{Var}_1\left(\hat{\theta}_i \middle| 2\right)\right]$$

where

1 = sampling stage where $\hat{\theta}_i$ estimates θ_i,

2 = sampling stage where θ_i is a random sampling from the population values of θ,

then

$$\mathrm{Var}\left(\hat{\theta}_i\right) = \mathrm{Var}_2\left[\theta_i\right] + E_2\left[\mathrm{Var}\left(\hat{\theta}_i \middle| \theta_i\right)\right]$$

$$= \sigma^2_{\theta_i} + \overline{\mathrm{Var}\left(\hat{\theta}_i \middle| \theta_i\right)} \qquad (1)$$

and where

$\sigma^2_{\theta_i}$ = natural variance in θ_i,

$\overline{\mathrm{Var}\left(\hat{\theta}_i \middle| \theta_i\right)}$ = average measurement error.

Using Eq. (1) and the method-of-moments, where

$$E\left(s^2_{\hat{\theta}_i}\right) = \sigma^2_{\theta_i} + \overline{\mathrm{Var}\left(\hat{\theta}_i \middle| \theta_i\right)},$$

then an estimate of natural variability can be calculated as follows:

$$\hat{\sigma}^2_{\theta_i} = s^2_{\hat{\theta}_i} - \overline{\hat{\mathrm{V}}\mathrm{ar}\left(\hat{\theta}_i \middle| \theta_i\right)} \qquad (2)$$

where

$$s^2_{\hat{\theta}_i} = \frac{\sum_{i=1}^{n}\left(\hat{\theta}_i - \bar{\hat{\theta}}\right)^2}{n-1}, \qquad (3)$$

$$\overline{\hat{\mathrm{V}}\mathrm{ar}\left(\hat{\theta}_i \middle| \theta_i\right)} = \frac{\sum_{i=1}^{n}\hat{\mathrm{V}}\mathrm{ar}\left(\hat{\theta}_i \middle| \theta_i\right)}{n}, \qquad (4)$$

$$\bar{\hat{\theta}} = \frac{\sum_{i=1}^{n}\hat{\theta}_i}{n}.$$

In other words, you can estimate the natural variance in responses $\left(\sigma^2_{\theta_i}\right)$ such as SARs, TIRs, or D based on the empirical variance among the replicate values [Eq. (3)] and average measurement error [Eq. (4)] without any distributional assumptions whatsoever. The only assumptions are:

1. $\hat{\theta}_i$ is an unbiased estimator of θ_i.

2. $\hat{\text{V}}\text{ar}\left(\hat{\theta}_i \middle| \theta_i\right)$ is an unbiased estimator of sampling error.

3. A random sample of the population of inference.

In the case where seasonal trends exist as indicated in travel times (Figs. 2,3-2.8), regression can be used to describe the pattern, leaving the error mean square (MSE) as an estimate of total variability [Eq. (3)]. This MSE can then be partitioned into natural variation about the trend and measurement error.

- There are several implications of the parametric approach taken to variance component estimation used in the CSS report, including the following:

1. Incorrectly using a binomial variance for the measurement error of the SARs will underestimate that component and overestimate natural variation $\left(\sigma^2\right)$.

2. The CSS report neglects to present the exact form of the beta distribution used, and there is an entire family of beta distributions to choose from. In the typical beta distribution, the means and variances are as follows:

$$\mu = \frac{\alpha}{\left(\alpha + \beta\right)} \qquad (5)$$

with a variance of

$$\sigma^2 = \frac{\alpha\beta\left(\alpha + \beta + 1\right)}{\left(a + \beta\right)^2}. \qquad (6)$$

If the CSS approach is correct, then the values $\alpha/\left(\alpha + \beta\right)$ for the SARs in Table 4.1 should be very close to the average SAR values across years. Unfortunately, the exact parameter estimates used in their calculations is not provided in the report. Such critical information and evaluation of assumptions need to be included in this report.

3. Similarly, if the fitted beta distributions are adequate, the beta variance (6) should reasonably approximate the nonparametric estimates of Eq. (2) and should be compared. Again, this critical information and analysis are not presented in this report.

4. The assumptions that SARs are beta-distributed are critical to the inference concerning the frequency of events. A goodness-of-fit test to the beta distribution needs to be performed using, for example, a Kolmogorov-Smirnov test, to verify the assumptions.

5. The use of the beta distribution to describe the frequency of SARs assumes the observed data are independent and identically distributed. However, this contradicts the results in Chapter 2, where the inriver survival, which contributes to the overall SARs, was found to be significantly correlated with environmental factors (e.g., Table 2.1). In other words, annual conditions influence the values of SARs for different stocks. The beta distribution ignores that previous set of findings and ignores the expected distribution of environmental conditions in the past or possible future. This should include projecting the possible consequences of global warming on inriver conditions and subsequent SARs.

Page 88, Table 4.1

- The estimated demographic variance is greater than total variance, suggesting something is wrong and thus casting doubt on all methods and results in this chapter.

- Observed correlations between point estimates of SAR for transport and C_0 groups for wild steelhead are explained by small transport groups and so are not used. However, such small transport groups (we are not told the actual sizes) produce unreliable parameter estimates that can seriously distort interpretation of results.

Chapter 5 Review

Chapter 5 compares estimates of annual SARs to target values indicated by the Northwest Power and Conservation Council (NPCC) (2003), and historical SARs based on run reconstruction methods. Multiple regressions are reported, relating Chinook salmon SAR to environmental variables. Upstream-downstream comparisons are made between Snake River Chinook and Chinook salmon from the John Day River. Biological comparisons between Snake River and John Day River Chinook are reported.

Introduction (pp. 105-106)

Page 106, lines 11-22

- Critiques of the single release-recapture (SR) analysis and PATH have demonstrated the reliance of latent mortality results on untestable assumptions, e.g., stock-specific Ricker *a*'s versus a common Ricker *a*. Additionally, climate effects have been shown to account for the majority of latent mortality. These criticisms should be addressed in this chapter.

Page 106, lines 19-20

- It is not explained and it is unclear how direct mortality, differential delayed mortality of transported smolts, and the common year effect were accounted for in the SR comparisons.

Methods: General (pp. 107-109)
Page 107, line 26

- "Overall SAR" is never defined, either here or elsewhere in the report. Presumably it is equal to "annual SAR," which is also never defined analytically.

Methods: Relationships between Chinook SAR and environmental covariates (pp. 110 - 115)

Pages 110-111

- Run Reconstruction SARs: Include jacks and adults; measure returns to mouth of Columbia River.

- CSS SARs: Include only adults (Chinook), no jacks; Measure returns to LGR

- Are run-reconstruction SARs and CSS SARs really comparable? It has not been justified that direct comparison of the measures is appropriate.

Page 112, lines 15-19: How is WTT defined?

Page 114, lines 3, 9; Figures 5.5, 5.6

- What does "frequently incorporated in multiple regression models" mean?

Page 115, Multiple Factor Model, lines 5-31

- How were candidates for independent environmental covariates selected? What were they? Only WTT, PDO, and an upwelling index are named, and it is unclear whether other covariates were considered.

- Harvest and temperature are known to affect SARs and do not appear to have been considered.

- Were any other “inriver” predictors than WTT considered?

- Were interaction terms considered in the multiple regression models?

- Typo in SAS version (presumably 9.1, not 91).

Methods: Snake River and Downriver SAR Comparison (pp. 115-119)

- There has been much previous criticism of the upriver-downriver comparisons made by the CSS and of the spawner-recruit model used to justify the upriver-downriver comparisons. Insufficient response has been made to these criticisms.

Page 116

- Lines 7-8

  - How is μ_t defined and estimated? Provide an equation showing how value is calculated. Is this the same μ as in Eq. 5.3, or is it the differential mortality defined verbally based on Eq. 5.2?

  - The “delta model” should be defined.

- Equation 5.3: If there is no delayed mortality from hydrosystem, then we expect $\exp\left(-\mu_{SAR,t}\right) = S_{J(\text{LGR-JD})}$.

  This important point is omitted from the report.

Page 117

- Line 18: Only a single hatchery (Carson) is used for the downstream hatchery Chinook salmon.

Page 119, Table 5.9

- This table is very difficult to understand. The caption does not agree with the notation used in the table. Values reported in the table are not sufficiently explained. It appears that the formula used to estimate BOA detection efficiency (*p*) is wrong.

- What are the values reported in the row “GRA, MCA, IHA?”

- MCN and IHA are not mentioned in table caption.

- The estimate of *p* based on detections at BON and upstream is INVALID if it is based on detections from different years, unless upriver adult survival to GRA is constant across return years, and detection probabilities at MCN, IH, and GRA are constant across return years. This is not true, so estimates of BOA detection efficiency presented here are invalid.

- On a yearly basis, *p* should be estimated as,

$$\hat{\rho} = \frac{\text{No. detected at BON and upstream}}{\text{No. detected upstream of BON}}$$

(from Manly-Parr) or equivalently as

$$\hat{\rho} = \frac{\text{No. detected at BON}}{\left(\text{No. det. at BON}\right) + \left(\text{No. det. upstream, not at BON}\right) \times \left(\text{Survival from BON to upstream}\right)}$$

(from CJS model), but NOT as is estimated here:

$$\rho_{CSS} = \frac{\text{No. detected at BON}}{\left(\text{No. det. at BON}\right) + \left(\text{No. det. upstream, not at BON}\right)}$$

The estimates of *p* as reported by the CSS will be positively biased, i.e., too large.

Methods: Comparison of biological characteristics of Snake River and downriver smolts (pp. 119-121)

Page 120

- In general for upstream/downstream comparisons, was goodness-of-fit considered or examination of residuals performed? Show results.
- With only 6 years of data, this is not a long time series, which limits the amount of useful information that can be gleaned from it.

Page 121

- Lines 13-14: No migration distance is given for JDAR1 fish. Comparison of survival and travel time between upstream and downstream fish should incorporate migration distance for the two groups of fish.
- Lines 40-41: Basing analyses on (Number of BON detects/Number released at trap) assumes that all groups have the same conditional detection probability at BON. This is likely to change with arrival timing.

Results: Overall SARs (pp. 122-127)

Page 122, lines 32-34

- "Removing sampling variability" resulted in lower mean SAR. Does this always occur?

Page 126, lines 17-19

- The CSS has been using a geometric mean previously, but here does not identify the type of mean used for mean SAR.
- It is unclear what the reference to the *t*-distribution means. If a formal *t*-test is being performed, this should be stated simply. Note that while these arithmetic means may be compared using a *t*-distribution, the geometric mean should not.

Results: Relationship between SAR and environmental covariates (pp. 128-131)

Page 128

- Lines 4-8: The data for the PIT-tag SARs and environmental factors are not presented in this report.

- In general, references to figures should be proofread. There are mistakes in figure references throughout the chapter, making it difficult to follow the narrative.

- Line 11: What is meant by "bi-variate results?" Is this regression of a single response variable on a single predictor variable? A vector response variable on one or more predictor variables? A single response variable on two or more predictor variables?

- Table 5.4: Did the CSS consider correlation between PDO and UP45n? Both types of measures are used in the same regressions, apparently.

Page 129, Table 5.5

- It should be explained why SepPDO is used rather than JulyPDO as a covariate, when JulyPDO looks better than SepPDO for both the long and current time series.

Page 130

- In general for regression with environmental variables:
  - What was the set of candidate predictor variables? Was it only PDOs, UP45ns, and WTT?
  - How model selection was performed needs to be specified?

- Lines 13-14 – The report says that WTT was "less significant for the shorter time series," but Table 5.7 indicates that WTT was not at all significant if the model includes upwelling index (Table 5.7, Current Time Series).

Page 131, lines 1-2, and Table 5.7, Current time series

- What model selection criterion were used to identify the "best" model?

- The "best" model shows no predictor variable significant at the 10% level when upwelling index is included.

- Without upwelling index (NovUP45n), SNWTT and MayPDO become significant. Was multicollinearity between these parameters and NovUP45n considered? And how?

<u>Results: Snake River and Downriver SAR Comparisons (pp. 131-136)</u>

- The CSS upstream-downstream comparison of SARs is based on the performance measure $U/D = \frac{S_{LGR-BON}}{S_{JD-BON}}$. If there is no differential post-JD mortality for upstream fish, then we expect U/D to equal S_{LGR-JD}, inriver smolt survival from LGR to JD.

- The CSS also reports values of $\mu_{SAR} = -\ln\left(U/D\right)$ for wild Chinook, although not for hatchery Chinook salmon. There is no benchmark for μ_{SAR}, however, because it compares the SAR from LGR to BON for upriver stocks to the SAR from JD to BON for downriver stocks.

- Interpretation of both U/D and μ_{SAR} estimates depends on the inriver survival of upriver stocks from LGR to JD, which is never considered by the author.

- Tables 5.8 (p. 132), 5.9 (p. 133), and 5.10 (p. 135), and Figure 5.16 (p. 136) cannot be usefully interpreted as they are, because they do not compare the reported measures to S_{LGR-JD}. Figure 5.16, showing the pattern of μ_{SAR} across years for wild and hatchery Chinook salmon, demonstrates the variation in μ_{SAR} across stock. Without also showing S_{LGR-JD} across stock, however, it is impossible to reach any conclusions.

- Table 1 (below) shows CSS estimates of U/D and μ_{SAR} taken from Tables 5.8-5.10, and compares them to estimates of S_{LGR-JD} calculated from Tables D-31 through D-36. Using the criterion of $S_{LGR-JD} > U/D$, or equivalently, $\exp\left[-\mu_{SAR}\right] < S_{LGR-JD}$, it is determined whether upstream stocks had lower SARs from JD to BON (SAR_{JD-BON}) than downstream stocks (referred to as "differential mortality" by the CSS).

  - In 4 of 5 years, wild Chinook upstream stocks showed lower SAR_{JD-BON} than wild Chinook stocks from John Day River (i.e., differential mortality).

  - Rapid River Hatchery spring Chinook salmon showed no differential mortality.

  - Dworshak Hatchery spring Chinook salmon showed differential mortality in 3 of 5 years.

  - Catherine Creek Acclimation Pond spring Chinook salmon showed differential mortality in only 1 of 4 years.

  - McCall Hatchery summer Chinook salmon showed no differential mortality in 5 years.

  - Imnaha Acclimation Pond summer Chinook salmon showed differential mortality in only 1 of 5 years.

  - In some years, the U/D measure is considerably greater than S_{LGR-JD}, such as Rapid River spring Chinook salmon for 2003, when U/D was estimated at 1.21 and S_{LGR-JD} was estimated at 0.502. There are similar examples from most hatchery Chinook stocks, in which U/D is estimated to be greater than 1, and S_{LGR-JD} is estimated to be less than 1. The CSS report does not address this situation, and gives no indication how U/D > 1 should be interpreted. In these cases, upstream stocks had higher SARs to BON, whether from LGR or from JD. If we were to follow the CSS's example, we must conclude that passage through the hydrosystem *improves* survival for many upstream hatchery stocks.

– It is obvious from Table 1 that the value of U/D (and by extension, μ_{SAR}) alone does not indicate whether or not "differential mortality" has occurred.

Table 1. Comparison of $U/D = (SAR_U)/(SAR_D)$ from Tables 5.8 (p. 132) and 5.10 in the CSS report to estimated inriver smolt survival from LGR to JD, calculated from Tables D-31 to D-36. If S_{LGR-JD} is greater than U/D, then upstream fish had a lower SAR than downstream fish from JD to BON.

Stock	Year	$\frac{SAR_U}{SAR_D}$	S_{LGR-JD}	SAR (JD-BON) lower for upstream stock?
Wild Chinook	2000	0.24	0.622	Yes
	2001	0.47	0.377	No
	2002	0.31	0.704	Yes
	2003	0.12	0.693	Yes
	2004	0.15	0.542	Yes
RAPH Sp Chinook	2000	0.79	0.741	No
	2001	0.76	0.529	No
	2002	0.83	0.745	No
	2003	1.21	0.502	No
	2004	0.50	0.508	No
DWOR Sp Chinook	2000	0.46	0.658	Yes
	2001	0.24	0.380	Yes
	2002	0.59	0.676	Yes
	2003	1.11	0.683	No
	2004	0.63	0.583	No
CATH Sp Chinook	2001	0.20	0.389	Yes
	2002	0.87	0.721	No
	2003	1.25	0.694	No
	2004	0.66	0.570	No
MCCA Su Chinook	2000	1.09	1.07	No
	2001	0.81	0.399	No
	2002	1.35	0.840	No
	2003	2.85	0.749	No
	2004	0.69	0.627	No
IMNA Su Chinook	2000	1.05	0.655	No
	2001	0.45	0.547	Yes
	2002	0.73	0.640	No
	2003	2.50	0.765	No
	2004	0.78	0.642	No

Page 132, Table 5.8

- How are the SARs for downriver wild Chinook salmon estimated? If simple return ratio, why not use same method for Snake River fish?

Page 134

- Lines 2-5: The CSS claims that the SAR to BON is *always* higher for the downriver (hatchery) fish, but that is not true for 2003.

- Lines 13-16: The reason given for not providing a confidence interval on SAR for downriver fish in 2004 is because an average survival to BON from previous years is used. However, that survival is not known without error, so a measure of uncertainty should be reported on survival to BON for 2004, and that error could be propagated to produce a CI on SARs.

- In general, the CSS addresses uncertainties incorporated by using a single downstream hatchery stock when the upstream/downstream results show no effect of the hydrosystem (i.e., for hatchery Chinook salmon), but not when the upstream/downstream results *do* imply hydrosystem effects (i.e., for wild Chinook salmon). This sounds like an inconsistent approach.

Page 135, Table 5.10

- In some years, upriver SAR > downriver SAR for hatchery Chinook salmon, despite additional inriver migration for upriver fish. Presumably, this result is unexpected and should be addressed. Such results may be due to large measurement error that obscures the relationship or the upstream/downstream pairing is a mismatch.

<u>Results: Comparison of Biological Characteristics of Snake River and downriver smolts (pp. 136-143)</u>

Page 136, Figure 5.16 - Needs confidence intervals.

Page 139

- Lines 2-3 It says that there is a significant ($P < 0.001$) difference in density-adjusted mean fork lengths of 106 and 106 mm (for IMNTRP and JDAR1), and separately of 100 and 100 mm (for SALTRP and SNKTRP).

- Lines 6-7**:** The report is inconsistent when it says 74 mm vs. 121 mm in fork length is not significant, especially considering that they previously defined any differences >5 mm to be biologically significant.

Page 141, lines 11-13

- "Smolts from upriver populations and downriver-origin smolts migrated at a similar rate, once their different migration distances were accounted for." What does this mean? Their migration "rate" (i.e., distance traveled per unit time) already accounts for differing migration distances.

Page 142

- Lines 15-17 – The observation that upriver smolts took longer to travel to BON than downriver smolts is not surprising since they leave at the same time and travel at the same rate, given that upriver smolts have farther to travel.

Results: SARs by Bonneville Arrival Timing (pp. 144-146)

Page 143, lines 1-2

- The "pattern of delayed arrival" was *not* consistent across years, as is stated – See years 2000 and 2003.

Page 144

- Lines 11-14 – What reference point is used to determine that upstream smolts experience delayed migration?

- Lines 16-18 – What does "significantly experienced lower SARs" mean? Does this mean that the difference or ratio between the SAR for wild upstream Chinook and wild downstream Chinook was *statistically* significant? *Biologically* significant?

Page 145, Figure 5.23

- Binomial confidence intervals are shown, but error is not binomial for C0, C1, and T0. Recalculate appropriately.

- In some years, large numbers of upriver migrants are omitted from the analysis by restricting attention to 16 April – 31 May window.

PIT-tag SARs versus SAR of run-at-large (p. 147)

- Lines 3-5: Are run-reconstruction SARs and CSS SARs mathematically comparable? Justify.

- Lines 12-19: Assumptions necessary for the run-reconstruction SARs are discussed, but not assumptions for the CSS SARs.

Discussion (pp. 148-151)

Page 148

- Lines 21-22: The limitations of small sample size cannot be avoided by using multi-year methods, as indicated here. Multi-year methods result in conclusions that are based on many uncertain estimates (due to small yearly sample sizes), instead of based on only a single uncertain estimate. This simply expands the problem of small sample size.

- Lines 29-31: WTT is named the "best" predictor variable for SARs, but it is not clear that the CSS considered other inriver covariates.

- Lines 37-38: It was found here that WTT influences the smolt migration rate. But JDAR1 and Snake fish have similar migration rates. Did they have different WTT? This needs to be addressed.

- Lines 42-43: SARs of downriver fish are compared to SARs from upriver fish, but these SARs are estimated over DIFFERENT reaches and distances, so we *expect* them to be different. The CSS needs to investigate whether the differences are more than expected.

Page 149

- Lines 39-40: "Hydrosystem migration rates did not differ between groups but were strongly influenced by water travel time." It is not clear how to interpret this statement. Did groups have different water travel times but the same migration rate? Or did they have the same WTT? Or was migration rate and travel time examined on an individual fish basis, instead of a group basis? It is not clear.

- Lines 41-46: Distance to travel is not considered as a factor of travel time.

Page 150

- lines 1-6 – It is claimed that the "potential confounding effects due to life history differences are probably negligible," but the CSS does not attempt to model SAR using both the upstream/downstream designation and the life history differences. Additionally, the effect of distance to travel was ignored. A model that includes all possible factors affecting SAR should be considered, in order to claim that it is the hydrosystem rather than other factors that cause the difference in return rates.

<u>In general for Chapter 5</u>

- In order to determine if there is a biological difference that explains any differences in SAR between upriver and downriver stocks, model SAR using fork length, migration date, arrival timing, year, in addition to upstream/downstream classification. Is upstream/downstream effect significant, given presence of all others?

- Looking at population differences in fork length, migration date, etc., one at a time, is reasonable for initial data exploration, but insufficient for conclusions about the significance of the upstream/downstream effect.

<u>Throughout Chapter 5</u>

- Typos are made in references to tables and figures throughout the entirety of Chapter 5.

- Pages 139–144: The reader is referred to a nonexistent figure for release site abbreviations.

Chapter 6 Review

Page 154, line 9

- The notation RY has not been defined. The context suggests Return Year, but Release Year is also a possibility.

Page 154, Tables 6.1 and 6.2

- Pooling migration success data across migration year and return year is valid only if those factors are nonsignificant. Perform test of homogeneity.

- Also applies to Page 155 (lines 17-23); Page 156 (Table 6.3).

Page 155, line 41

- Was return year modeled as a fixed or random effect? Most blocking factors are modeled as random effects, although there are times when a fixed effect is more appropriate.

Page 156, Table 6.3

- Chi-squared tests indicate whether there is a difference in perceived upriver adult survival across juvenile migration groups, but they do not indicate the nature of the difference. The p-values reported do not indicate that the actual ranking in the Success Rate Ranking column is significant, simply that at least one of the juvenile migration groups had a significantly different success (survival) rate than the others. One-sided tests should be performed comparing pairs of juvenile migration groups in order to test the significance of the ranking.

Page 157, Figure 6.1

- Needs error bars or confidence intervals.

Page 158, Figure 6.2

- Needs error bars or confidence intervals.

- The interpretation of Fig. 6.2, showing the proportion of LGR-detected adults and jacks detected at hatcheries, depends on the detection effort at each hatchery in each year. Without that information, it is useful only for comparing transported to inriver fish. It appears that transported fish had slightly better survival from LGR to the hatcheries, but without error bars and without information about detection effort (and harvest pressures, etc.), no real conclusion can be reached from Fig. 6.2.

Page 159, lines 11-15

- The overall average perceived BON-LGR adult survivals for the three migration groups are not very useful without standard errors or confidence intervals.

- It is not clear how these average survivals were computed. Were yearly estimates weighted by the number of fish returning in each year? Or were migration year estimates averaged?

- Given the finding that return year is a significant factor in perceived upriver adult survival (from the logistic regressions presented later in this chapter), pooling data over return year is not warranted.

Page 160

- Lines 16-19: The model evidence ratio does not indicate that one model is “more likely” than another, in a Bayesian sense. Rather, it means that there is more evidence for one model compared to the others.

  - Also applies to results for wild Chinook salmon (p. 161, lines 22-23).

  - The highest evidence ratio for the best model for wild Chinook salmon (p. 161, lines 21-25; p. 162, Table 6.6) is at most 4, thus there is *not* clear evidence that transportation is an important factor in determining adult migration success when compared to environmental factors.

- Lines 29, 32: It is not clear how the confidence intervals on the odds ratios are computed. Provide explanation. Asymptotic normal-theory confidence intervals are considerably narrower than those reported, and do not include 1 for either LGR-transport fish or LGSdown fish. If the confidence intervals were based on a *t*-distribution, the degrees of freedom should be reported (Table 6.5).

Page 161

- Table 6.5

  - Degrees of freedom should be reported for each parameter estimate.

  - Surprisingly, warmer temperatures were associated with higher perceived adult survival. Perhaps temperature is confounded with run (spring versus summer).

- Lines 29-30: The odds ratio is misinterpreted here. An odds ratio of 0.5 does *not* mean that the probability of success of LGR-transport fish is half that of inriver fish. If the probability of success (i.e., perceived adult survival from BON to LGR) is P_{LGR} for LGR-transport fish, and is $P_{inriver}$ for inriver fish, then:

$$\text{Odds ratio} = \frac{\left(\dfrac{P_{LGR}}{1-P_{LGR}}\right)}{\left(\dfrac{P_{inriver}}{1-P_{inriver}}\right)} = \frac{1}{2}$$

$$\Rightarrow P_{LGR} = \frac{P_{inriver}}{2-P_{inriver}}$$

This means that the probability of success of LGR-transport fish depends on the value of the success probability for inriver fish, as demonstrated in Table 2 below. Table 2 indicates that for an odds ratio of 0.5, the probability of success of LGR-transport fish is generally greater than half that of inriver fish, except for very small inriver success probabilities, which are not applicable here.

Table 2. The probability of adult migration success (BON to LGR) for inriver fish and LGR-transport fish for an odds ratio of 0.5.

$P_{inriver}$	P_{LGR}	$\frac{P_{LGR}}{P_{inriver}}$
0.1	0.05	0.53
0.25	0.14	0.57
0.33	0.2	0.6
0.5	0.33	0.67
0.75	0.6	0.8
0.9	0.82	0.91
1	1	1

Pages 162-163, Hatchery Chinook arrival and travel time ANOVAs

- For both arrival time and travel time, the interaction term between return year and juvenile migration (outmigration) method was significant. This affects interpretation of the main effects of both return year and outmigration method, so conclusions based on the main effects alone are invalid.

- The ANOVAs should be included in the report.

Page 164, lines 38-40

- How much of TIR or D is explained by observed differences in perceived upriver adult survival between inriver and LGR-transport fish?

Chapter 7 Review

Page 168; lines 31, 46

- Survival from release to LGR = 0.95 seems high, and does not correspond to year 2000 data used as basis for default values

- SAR = 0.03 seems high, given that observed SAR has been lower than the target value of 2% in most years according to this report.

Page 169, line 9

- Joint probability of survival from BON to TWX and detection at TWX = 0.10 is high, based on past years.

Page 170 (lines 9-11, 21-23, 39-41), Page 171 (lines 8-10)

- The survival probabilities used in simulation scenarios #5, 7, 10, and 12 include inriver survival probabilities > 1, when the variable *day* is 0 or very low. Inriver survival should be parameterized using only admissible parameter values (i.e., ≤ 1) and included in this report.

Page 171, last paragraph, and Page 174

- The emphasis on the T and R groups is confusing. The underlying cause of the assumption violation is not that R fish have higher or lower inriver survival than T fish, but that detected fish have higher or lower inriver survival than non-detected fish. While understanding the effect on the T and R groups will be useful in the future, it is not clear how they apply to estimation of C0, C1, and T0 for previous years' data, in which T and R groups were not used. A simpler method of assessing the effect of detection-influenced survival would be to simulate data under the scenario described (Post-turbine survival < Post-bypass survival < Post-spill survival, with varying proportions of undetected fish passing via turbine or spill) and examine estimates of C0, C1, and T0.

Page 175, last paragraph

- It is not clear if the "true" survival parameters used to compute LGR equivalents are averages of seasonal survival parameters, or if LGR equivalents are computed on a daily basis and then summed over the season. Given the temporal variation in survival parameters introduced in these simulations, the latter approach should give a better representation of the "true" C0, C1, and T0 groups. Clarify approach and, if necessary, rerun simulations.

Page 194

- The CSS uses results of the second set of simulations to address how to best analyze data using the NPT approach, in which tagged fish are pre-assigned into migration groups: T (transport) fish are transported upon their first detection at a transport dam; R (river) fish are returned to river upon all detections. Using the C0/C1/T0 approach to analyze data with pre-assigned migration groups is not intuitive. It would be simpler and more defensible to simply compare the SAR of the T group to the SAR of the R group. All "R" fish will have migrated wholly inriver, while some "T" fish will have been transported and others (undetected) will have migrated inriver. The comparison of SAR(T) to SAR(R) is more easily interpreted for management, because the alternative to transportation is to return detected fish to the river, whereas the transportation alternative being tested in the SAR(T0) vs. SAR(C0) comparison is not clear.

Chapter 8 Review

Page 198

- Lines 35-38: The trend of performance measures for wild fish mimicked the overall trend of performance measures for the collection of hatcheries, but did not agree well with the trend from any single hatchery across all years. It is not clear which single hatchery could be used to make inference to wild fish. Also applies to Pages 199-200, bullet (b) of Chapter 5 summary.

- Report confidence intervals for results (e.g., geometric means).

Page 199

- Lines 32-35: The inference made from declining SAR(C1) over the season to hydrosystem-caused post-Bonneville mortality is unfounded. There are alternative possible causes of post-Bonneville mortality, including temperature, pollution, disease, and seasonal changes in estuary conditions. No conclusions about the relative importance of the various potential sources of mortality can be reached here.

- Lines 40-42: The CSS claims that Snake River wild steelhead SARs averaged less than 2%. It is difficult to confirm this statement, because the annual SARs are not presented in tabular form in this report. However, Fig. 3.12 suggests that average annual SAR for wild steelhead may be greater than 2%. Document annual SARs in the table and explain apparent inconsistency.

Page 200

- Lines 8-14: The CSS did not compare the ratio of upstream and downstream SARs to inriver survival between Lower Granite and John Day, so the conclusion that upstream fish experience extra mortality caused by the hydrosystem is unjustified. Also applies to Page 200 (lines 33-34).

- Lines 23-24: The claim is made that that the CSS shows clear evidence of delayed estuary entry of Snake River inriver smolts, caused by passage through the hydrosystem, on the basis of comparisons with John Day smolts. This is not true. The CSS found that Snake River and John Day smolts (1) initiate migration at the same times, and (2) migrate at similar rates through the first three dams passed. Given the extra distance traveled by the Snake River smolts, it is not surprising that Snake River smolts enter the estuary later than John Day smolts. The CSS analysis would be more useful if it had compared the observed and expected arrival dates of the Snake River fish, given their migration initiation date, migration rate (through the first three dams), and distance to travel.

- Lines 26-30: The conclusion that differing seasonal SARs for upstream versus downstream smolts is evidence of delayed mortality ignores possible alternative explanations, including potentially different ocean residencies.

- Lines 37-42: It appears here that wild and hatchery Chinook salmon transported from LGR always had 10% lower SAR than fish passing through the hydrosystem by alternative routes. It should be noted that the effect for hatchery fish (4% to 7%) was considerably less than the effect for wild fish (15%), so the 10% effect reported is somewhat misleading.

Page 202

- Lines 11-16; lines 39-41: The claim is made that the CSS addresses the question of whether smolt transportation compensates for effects of the Federal Columbia Power System (FCRPS) on survival of Snake River chinook and steelhead. This claim extrapolates past the available data. The CSS compares the SAR of transport fish to the SAR of fish migrating inriver. While the inriver fish experience effects of migrating through the FCRPS, available data do not indicate the magnitude of those effects; this would require comparing the SAR of fish migrating through the FCRPS to the SAR of fish migrating through the same reaches but not through the FCRPS. That is not possible. At most, the comparison of the SAR of transport fish to the SAR of inriver fish indicates whether transportation is a viable management option; it is *not* equivalent to comparing transportation to migration through the unimpounded river. It is worth noting that the SAR from BON to BOA for hatchery Chinook salmon from the John Day river was less than 2% for 2001 through 2004 (Table 5.10). Regardless of the validity of upstream-downstream comparisons, these low SARs for John Day fish suggest that the hydrosystem is not the only factor in below-target SARs.

- 3rd paragraph:

  - The CSS reports "mean" values for TIR for steelhead, failing to mention that these are geometric means. Typically, "mean" implies the arithmetic mean. Geometric means produce lower values than arithmetic means. Omitting 2001, the arithmetic mean of TIR for wild steelhead was 2.4, versus a geometric mean of 1.7; the arithmetic mean TIR for hatchery steelhead was 1.7, versus a geometric mean of 1.5.

  - The CSS says that D < 1.0 for steelhead. However, wild steelhead showed D values >1.0 in 5 of 7 years, with an arithmetic mean of 1.12 (including 2001). Thus, it appears that in most cases, D > 1.0 for steelhead. This inconsistency should be explained.

  - When TIR values are at 1.0 or greater, the CSS points out that D values are nevertheless less than 1.0. They do not discuss the implications of this. Even if D < 1, the decision to transport should be based on TIR values, not on D.

- Last paragraph: The CSS compares observed SAR estimates from PIT-tagged fish to the NPCC objectives for SAR (2% minimum, 4% average), without addressing the NOAA finding that PIT-tagged fish have lower survival than untagged fish (as requested by the ISAB). Without knowing the size of the PIT-tag bias, comparisons of PIT-tag SAR to target values are not completely useful.

Page 203, 3rd paragraph

- The CSS mentions that the decision of when to initiate transportation is an important management decision, and implies that this study fully addresses that question. While some estimation and analysis of seasonal TIR was done, it was hardly a complete analysis, and provides little management guidance.

- The CSS claims that seasonal TIRs "may contain some positive bias" because they are based on the C1 group (detected, not transported) rather than the C0 group (undetected). However, because the management alternative to transportation is to return bypassed fish to the river, the C1 group is more appropriate than the C0 group for comparison to transport SARs.

Appendix B Review

Page B-3, Figure 1

- The estimators of ϕ_1, ϕ_2, and ϕ_3 are correct.

- The figure is somewhat cryptic. The parameters ϕ_i are not defined, nor are the statistics R_i, $R'_{1\bullet 2}$, $R'_{12\bullet 3}$, …, r_i, m_i. The reduced *m*-matrix is not so standard that the CSS should expect all readers to recognize and understand it without further explanation. Provide more detail.

Page B-4

- The CSS explains that they allow individual reach survival estimates exceeding 100% when computing an overall multi-reach survival estimate. Why, then, do they not allow $\hat{S}_{JD-BON} > 1$ for 2004 for Carson NFH Chinook, in Chapter 5?

- The verbal description of the weighted average of survival estimates provided in the second full paragraph is insufficient. An equation demonstrating precisely how the overall survival estimate was estimated is required.

- The CSS used weights equal to the inverse relative variance of the reach- and cohort-specific survival estimates. How were the variances of those reach and cohort survival estimates computed? How was the standard error on the weighted average survival computed? Provide details.

- In the final partial paragraph, the CSS discusses using a "per-mile" expansion of juvenile survival in cases where it was impossible to estimate survival to BON directly. Previously (Chapter 3), they used a per-km method of extrapolation. Either there or here, did they consider any other basis for extrapolation? Did they consider the goodness-of-fit of the extrapolation method used? Did they estimate the standard error on the survival estimate to BON, either with or without the extrapolation?

Page B-5

- The CSS lists the three ways in which fish can pass an individual transport dam, and indicates that these three passage routes describe the passage routes through the hydrosystem. However, their three passage routes must be combined over multiple dams to describe the possible passage routes through the entire hydrosystem. For example, there are seven possible passage routes through LGR, LGS, and LMN that result in transportation from one of those dams.:

  - Transportation at LGR (route 1)

  - Transportation at LGS following either detection at LGR, or non-detection at LGR (routes 2 and 3)

  - Transportation at LMN following either detection at both LGR and LGS, detection at only one of LGR and LGS, or non-detection at both LGR and LGS (routes 4-7)

  Thus, the CSS "partition" of PIT-tagged smolts arriving at LGR is, at best, unclear from their description and, at worst, potentially omitting considerable numbers of fish. Clarification in this report is required.

Page B-6

- #5. Is “run-at-large” equal to “untagged” here, or does it also include tagged fish? How is P_{t2} estimated? Is P_{t2} really the proportion of the entire run-at-large that were transported at LGR, or only the proportion of the run-at-large collected at LGR that were transported? Similar comments pertain to #7 and #9.

Page B-7

- #13 - #15: It is essential for the CSS to actually write out the expected values of the statistics T0, T0*, and C1 in terms of the underlying model (i.e., survival, detection, transportation, and removal parameters), rather than leaving them partially defined. This level of technical detail is essential for all readers to know exactly what is being estimated by the parameters in the report.

- #15, #16: The statistics d0 and d1 are never defined. The 50% survival probability is not explained—50% survival to where? On what basis is 50% chosen? Why not use the actual estimated survival probability to whatever site or sites are used?

Page B-8

- Finally, an attempt is made at an explanation for the 50% survival probability used to deal with downstream removals. It is not sufficient, however. Why not use a dam-specific adjustment, rather than pooling all downstream removals and assuming a common survival to every downstream dam? Have the effects of violations of this 50% survival assumption been examined? It is known that violations of this assumption occur, because survival between MCN, JD, and BON is not 100%, so survival to one dam (e.g., at 50%) is not equivalent to survival to the other dams, as is implicitly assumed by using a single survival probability to all downstream dams. Additionally, if using a single survival rate is warranted and if survival to BON is to be used each year, it should be possible to use the estimated survival to BON for the year, rather than assuming 50% survival each year.

- Show the number of removals on a dam-specific basis that you contribute to d0 and d1?

- “Estimation of SARs for study categories:” SAR1(T0) and SAR2(T0) have been discussed but not defined in this report. Define all measures.

Page B-9

- A “common annual routing rate to the raceways” was used—what is this? Is it known or estimated? What value was used?

- The notation used to define AC0 and AC1 is insufficient. It does not preclude using adults that were removed at downstream dams for any reason. Because many removed fish are not sacrificed, it is conceivable that some of these “removed” fish may return as adults. Are these adults included in AC0 and AC1? One assumes not, because this would positively bias the SAR for the C0 and C1 groups; however, the notation used implies that these removed fish are included in AC0 and AC1.

- It looks like SAR2(T0) is used in this report for overall SAR of transported fish, rather than SAR1(T0), unless otherwise specified. Is this correct? Clarify.

Page B-11

- The expected value of the size of the C1* group should be presented. At the least, the definition of the C1* group should be explained. It does not make intuitive sense to define it in terms of the T0, C1, and T0* statistics, because the T0 and T0* statistics are based on different groups of fish.

Page B-12

- The CSS states that “the rate of harvest is assumed independent of whether fish had been transported as smolts. [These] assumptions … apply to both TIR and *D*.” Where does the CSS actually make use of this assumption? Is it only in their interpretation of results about TIR and D?

Literature Cited

Akcakaya, H.R. 2002. Estimating the variance of the survival rates and fecundities. Animal Conservation **5**: 333-336.
Burnham, K.P., and Anderson, D.R. 2002. Model selection and multimodel inference. Springer-Verlag, New York, New York, USA.
Burnham, K.P., Anderson, D.R., White, G.C., Brownie, C., and Pollock, K.H. 1987. Design and analysis methods for fish survival experiments based on release-recapture. American Fisheries Society Monographs **5**.
Gould, W.R., and Nichols, J.D. 1998. Estimation of temporal variability of survival in animal populations. Ecology **79**(7): 2531-2538.
Lowther, A.B. 2002. Development, expansion, and evaluation of release-recapture survival models for Snake River juvenile salmonids, with new algorithms allowing time-dependent individual covariates Dissertation, Quantitative Ecology and Resource Management, University of Washington, Seattle, Washington.
Neter, J., Kutner, M.H., Nachtsheim, C.J., and Wasserman, W. 1996. Applied linear statistical models. Irwin, Chicago, Illinois, USA.
Seber, G.A.F. 1982. The estimation of animal abundance. MacMillan, New York, New York.

CSS Response to ISAB Recommendations

1. **Describe methods clearly.**—Methods used to define and estimate the C0, C1, and T0 study groups and SAR, TIR, and D are presented. Methods used to define and estimate annual SAR are not presented clearly. Other methods (e.g., to remove sampling variability) are not presented fully or clearly.

2. **Report size at tagging to survival and relate to survival.**—Size at tagging is analyzed for the upstream-downstream comparison but is not reported for releases in general or related to survival.

3. **Address validity of inference from tagged fish to untagged fish.**—This point was addressed briefly, with criticisms of the methods used to determine that untagged fish have different survival than tagged fish. In general, results from tagged fish are compared to target values with no mention of any possible bias based on tagged fish.

4. **Use more downstream hatcheries in the upstream-downstream comparison.**—This was not done for the retrospective analysis.

5. **Do not limit analyses to an annual time scale; consider environmental and operational factors.**—Within-year patterns of SAR, TIR, and D are addressed briefly. The main focus of the analysis is on the annual time scale, due both to sample size and to the use of the C0 group, which cannot be analyzed on a smaller time scale. SAR, TIR, and D are related to several environmental factors. Operational factors are not considered.

6. **Perform a 10-year summary report.**—This is it.

7. **Test assumptions.**—Estimation results are analyzed for robustness to CJS-assumption violations. Little attention is given to whether or not those assumptions are violated.

8. **Pre-assign routes of passage to simplify analytical methods.**—This could not be done for the retrospective report. The simulations testing the robustness of estimation methods to CJS assumption violations incorporated pre-assigned routes of passage for future analysis. However, it appears that the analysis methods to be used with pre-assigned passage routes will remain unchanged, so the pre-assigned routes will not simplify analytical methods.

9. **Use more diverse metrics of differential survival (not only TIR and D).**—"Hydrosystem survival" was defined, but (1) was not used because it cannot be estimated, and (2) does not appear to be an improvement over TIR and D. "System survival" was used, but no expected or target values were given, and there was no

guidance for interpreting results. Conclusions continue to be based on TIR and D. Distance from ocean and hatchery practices were not considered.

General Technical Comments on CSS 10-Year Report

The Comparative Survival Study (CSS) 10-Year Retrospective Analyses Report provides a history of PIT-tagged salmonid fish performance from 1998 through 2006. The length of time, the breadth of geographic coverage, and range of salmon life-history phases investigated in the report have the potential of providing a valuable chronicle of recent Columbia River trends. No other study in the Fish and Wildlife Program has the same scope of effort.

Given the unique range of this project and the importance of this 10-year review, it is therefore unfortunate that the report does not document the tagging results more thoroughly.

- Reporting the empirical results of the tagging study is largely limited in this report to relatively few tables and graphs in Chapters 2 and 3. Appendix D supplements this information in these chapters but does not provide a showcase for the important trends and comparisons one might expect from a 10-year summary. For example, comparison of trends among the many hatchery stocks tagged is completely absent. Furthermore, in many cases, standard errors or confidence intervals for performance measures are neither reported nor displayed. The CSS Report needs to simply document and display the tagging results for the benefit of most readers and organizations that do not have the resources of the CSS organization. This issue is fundamental to our comment - the need to provide the means to reproduce results. (See also the closing comment on the last page.)

- Again we suggest that the CSS Report provide a straightforward presentation of tagging results. The Retrospective Report instead focuses on interpreting estimates of survival (S), smolt-to-adult ratios (SARs), transport-inriver ratios (TIR), and delayed mortality (D), using both previous as well as new approaches. This is unfortunate because it is in these analyses (as discussed in subsequent bullets) where the Retrospective Report most often falters in providing basic data and analyses useful for fish research and management.

- By definition, these PIT-tag studies are observational, thereby precluding direct causal inferences to any natural or anthropogenic factors. Replicated, randomized, and manipulative studies beyond the scope of the CSS study would be required for such inferences. Consequently, any attempt to identify environmental driving variables or differentiate ambient from hydrosystem effects is very difficult. The methods CSS uses in the report are not exempt from these problems, and contain several technical errors, as summarized below by chapter. Again, a direct causal inference to any natural or anthropogenic factor is precluded.

- Beyond that, there are conceptual issues, e.g., the approach of basing transportation analysis on C0 (undetected) fish. Not only does the C0 group not represent a real management alternative to transportation, but that group also migrates through the hydrosystem, not through an unimpounded river. Consequently, using the C0 group as a surrogate for a non-hydrosystem alternative is invalid. The CSS approach to estimating differential mortality using upstream-downstream comparisons is equally invalid. The CSS methods in many cases have not been peer-reviewed in the scientific literature, as might be expected for a 10-year-old program.

Below are summarized some of the more important technical concerns by chapter and also attached is a list of recommendations by the ISAB and our assessment of how well the CSS 10-year review complied.

- In Chapter 2, the concept of instantaneous mortality rate (Z) is introduced. However, it is not based on failure times (i.e., death times of PIT-tagged fish) as it properly should be, but rather on a simple function of the ratio of reach survival estimates and median travel times. The report inappropriately analyzes survivals, travel times (FTT), and Z as if they are three independent pieces of information. Reach survivals throughout the season are relatively stable while travel times show marked seasonal trends. Using the ratios of this information, Z values are calculated and seasonal trends are interpreted as survival processes. In fact, the trends in Z are essentially nothing more than the inverse trends in travel times construed as survival effects. Finally, the Report misinterpreted the instantaneous mortality rate, Z, as the probability of mortality (i.e., $1-S_t = 1-e^{-Zt}$), which it is not.

- Results of the modeling exercises on reach survivals in Chapter 2 are not discussed, leaving the impression that results for Z values are applicable to S, which is not true as discussed above

- The summary tables in Chapter 2 for the modeling exercise are difficult to interpret. Nevertheless, it appears as though models with partial regression coefficients (e.g., negative sign with flow) were routinely omitted. This practice left models that sometimes had an interaction term but no main effect, or a quadratic term without the linear component, which is inconsistent with general modeling-building practices. These nonstandard practices, as well as using Julian date as a surrogate for any number of unspecified environmental factors, greatly limited the interpretation and efficacy of the analyses.

- As requested by the Independent Scientific Advisory Board (ISAB), the CSS has compiled in Chapter 3 and Appendix B many of the methods used to generate the time series of estimates reported. Nevertheless, important definitions such as annual SARs, upriver adult survival, and project-specific TIR, are never mathematically defined in equations. Also missing are basic results, such as the numbers of fish actually transported at each dam, which should be documented in a 10-year review such as this report.

- The CSS report inferences are often based on whether confidence intervals overlap. Non-overlapping confidence intervals do not provide a valid test of significant differences. Instead, the CSS should use valid methods of testing significance, either within their bootstrap approach or separately with a parametric approach.

- Chapter 4 used parametric models to partition the total variance of SARs, TIR, and D into natural variation and measurement error. However, an underlying assumption, that the SARs are binomially distributed, is inconsistent with the mark-recapture models used to estimate the values. This invalid assumption results in underestimating the sampling error, which will inflate estimates of natural variation. The report needs to use goodness-of-fit tests to assess their parametric model assumptions and compare their parameter estimates with nonparametric variance components. Additionally, their inferences concerning natural variation do not take into account their own findings on ambient effects, the historical distribution of those ambient factors, or how influences such as global climate change may affect projections into the future.

- In Chapter 5, the CSS Report continues a practice of comparing upstream and downstream Chinook salmon stocks. As in the past, multiple upstream hatcheries and collection points are compared to only a single downstream hatchery and collection point (for wild fish), despite the ISAB's recommendation to incorporate more downstream stocks. Given that this is a retrospective report, it is understandable that the CSS report could not immediately include additional downstream stocks. However, the Report should have included the early data from downstream hatcheries that were originally used in the CSS. These hatcheries were removed from the study design by CSS management, contrary to the urging by some who viewed their inclusion as providing the exact perspective that the ISAB called for later in their 2005 review of the CSS.

- While the CSS does perform some useful comparisons of biological characteristics of the upstream and downstream stocks, their upstream-downstream analysis is invalid in other critical ways.

- The CSS uses an incorrectly conceived and constructed measure of "differential mortality"—just another name for latent mortality. The approach assumes that no natural mortality should occur for smolts between upstream and downstream sites. When the performance measure is corrected for the extra migration of the upstream stocks, there is little or no evidence of differential hydrosystem mortality for hatchery Chinook salmon.

- Additionally, in comparing travel times between upstream and downstream stocks in Chapter 5, the CSS report ignores the longer distance upstream stocks have to travel and then attributes their later estuary entry on the Federal Columbia River Power System (FCRPS).

- Finally, all these efforts in Chapter 5 to estimate latent mortality are contrary to the recent ISAB 2007 recommendations that such attempts be abandoned because this task is impossible with existing data. The report does not explain why it continues to pursue this rationale.

- Chapter 6 attempts to partition survival across different portions of the migration, focusing on smolt survival from the hatchery/trap to Lower Granite Dam (LGR), perceived adult survival from Bonneville Dam to LGR, and perceived adult survival from LGR back to the hatchery/spawning grounds. Adults are categorized by juvenile migration method. The effect of juvenile migration method—in particular, transportation—on adult upriver survival is an important question, and the analyses in this chapter relating adult survival to migration method are worthwhile. However, the CSS Report does not provide the methods used in estimating upriver survival for a given juvenile release group, so reviewing and reproducing their results is impossible. Additionally, the Report misinterprets the odds ratio from its own logistic regression when comparing adult survival for LGR-transport fish to other fish. Consequently, the Report overestimates the effect of LGR transportation on upriver adult survival.

- Chapter 7 describes the results of useful computer simulations to assess the effects of violations of key assumptions of the Cormack-Jolly-Seber (CJS) release-recapture model. The study demonstrated the robustness of the estimation methods to all but the most severe temporal changes in survival and detection probabilities. The results also showed the release-recapture model to be reasonably robust to changes in survival due to prior detection history. The focus of the simulation study was on the preassigned transport (T) and return-to-river (R) groups. Unfortunately, these two groups of fish do not directly translate into the C0, C1, and T0 groups used in the CSS analysis. Consequently, more focused simulations are still needed to assure the CSS methods are robust enough to model violations.

- The question of the effect of the FCRPS on salmonid migration and survival is important. However, it is not addressed by the analyses presented in this Report. Because the overall conclusions provided in the report are based on the invalid analysis of the previous chapters, the final inferences are unreliable.

- The CSS Report attributes all differences in survival and travel time between study groups in the upstream-downstream comparison to the hydrosystem, ignoring differences expected because of different migration distances and resulting natural mortality. Additionally, the Report incorrectly claims to have addressed the question of whether smolt transportation compensates for effects of the FCRPS on survival of Snake River Chinook salmon and steelhead. The comparison of the SARs of transport fish and inriver fish is not equivalent to comparing transportation to migration through the unimpounded river.

As urged by the ISAB, this CSS Report now presents some of the methods used in estimating SAR, TIR, and D for the various groups of interest (i.e., the C0, C1, and T0 groups). It has compiled methods strewn through previous reports in one place, and this makes the reading much easier. However, this encouraging start was not consistent across chapters. Our attempts to reconstruct final results from intermediate calculations presented in the report have been frustrated by a lack of necessary information and insufficient technical descriptions. The ability to reproduce results is crucial to the scientific peer review process. The Retrospective Analyses Report would have benefited from greater documentation of basic tagging results and from far less exploratory efforts to assign effects to the hydrosystem when causation really cannot be partitioned or identified using the CSS data. Due to the inability to reproduce these results using accepted modeling and analytical procedures the CSS Report's findings do not demonstrate the scientific rigor and support to authoritatively guide hydrosystem management.

FISH PASSAGE CENTER

1827 NE 44th Ave., Suite 240, Portland, OR 97213

Phone: (503) 230-4099 Fax: (503) 230-7559

http://www.fpc.org

e-mail us at fpcstaff@fpc.org

August 31, 2007

Dr Usha Varanasi, Ph.D.
Northwest Fisheries Science Center
NOAA Fisheries
Seattle, Washington

Dear Dr. Varanasi:

Thank you for your review of the Draft, Ten-year Retrospective Summary Report. The following response was developed by the Comparative Survival Study Oversight Committee, (Committee) comprised of, the Columbia River Inter-tribal Fish Commission, the Washington Department of Fish and Wildlife, the Oregon Department of Fish and Wildlife, the Idaho Department of Fish and Game, and the US Fish and Wildlife Service. As you are aware the Comparative Survival Study is a joint project of the agencies and tribes. The study design, the implementation of the study and the analysis are carried out collaboratively among the sponsoring fish and wildlife management agencies and tribes. The Committee has developed the following response to your general comments, which are followed by the response to each specific comment.

The CSS study uses regionally accepted analytical methodologies, and innovative approaches based upon peer-reviewed scientific literature. The methods and analysis are well within the methods and analytical approaches utilized by the Northwest Fisheries Science Center (NWFSC) in the 2005 Technical Memorandum available to the region. By working collaboratively on study implementation, design development and analysis, the experience and skills of the state, federal and tribal fishery mangers have been a valuable asset for this study. We have addressed the NWFSC comments on the CSS report in the attached (attachment 1) document.

The CSS Oversight Committee is grateful for the significant investment by NOAA in the review and preparation of comments on the draft report. The report has been improved as a result of addressing and incorporating comments. We look forward to future positive collaboration with NOAA on future CSS monitoring and evaluation.

Sincerely

Michele DeHart

Attachment 1

Reviewer Comment :*At the request of Paul Wagner and Ritchie Graves, we reviewed the DRAFT "Comparative Survival Study (CSS) of PIT-tagged Spring/Summer Chinook and Steelhead in the Columbia River Basin Ten-Year Retrospective Analyses Report." The report is extraordinarily long (377 pages); too long to read, digest and provide finely detailed commentary in the review time available. The following paragraphs summarize our major concerns with the report. Please call John Williams (206.860.3277) if you have any questions regarding these comments.*

Response: The main report is actually 212 pages (plus appendices), similar in length to the 2006 annual report. The NWFSC provided comments on previous annual reports. The ten-year report deadlines and the review schedule were determined by the NPCC, with little input from the authors. While we sympathize with the tight review schedule, we also note that the NPCC required schedule for report preparation was extremely tight for a ten-year report with this breadth and depth of analysis – November 2006 to June 2007.

Reviewer Comment:1. Most strikingly, despite its title and the fact that the CSS study group has PIT-tagged hundreds of thousands of juvenile Chinook salmon and steelhead, the CSS retrospective report does not contain a holistic analysis of this 10 –year effort or an integration of the results across all species that considers different migration conditions.

Response: The Merriam-Webster dictionary defines holistic as "relating to or concerned with wholes or with complete systems rather than with the analysis of, treatment of, or dissection into parts". The CSS Oversight Committee believes that we have presented and integrated the various components and analyses to present a holistic depiction of SARs and factors affecting SARs for the target species and study period, as requested by the ISAB. Certainly, with a large, robust data base such as provided by CSS, other analyses are possible and desirable.

This comment missed the substantial work that was done and presented throughout this report to holistically analyze the results that have been obtained to date through the CSS. Chapter 2 contains an extensive, holistic synthesis of observed fish travel time, survival and instantaneous mortality rates, along with an explicit evaluation of the effects of different migration conditions on these rates. The study covers a number of years for both species that reflect quite varied migration conditions (e.g., drought year 2001 versus high-flow year 1998). Further, within-season variation in SARs of both transported and in-river fish is explored in Chapter 4. In addition, we evaluated the influence of in-river, climatic and ocean conditions on Snake River SARs in Chapter 5.

Reviewer Comment: 2. The data presented and the discussion and conclusions' sections all seem focused through the lens of specific positions favored by the authors; hydropower system-related latent mortality is large in magnitude, transportation is not beneficial, management actions directed at the hydropower systems have generally failed, and consequently SARs have been low in recent years and drastic actions are needed to recover the wild Chinook salmon populations, as PIT-tagged wild fish fail to meet a minimum 2% SAR. Results that do not support desired *positions are usually discounted by carefully placed language. For example, from the conclusions in Chapter8 (all italics are ours):*

"Variation in [survival] in the MCN-BON reach was explained by temperature and Julian day. However, there was substantial uncertainty in the lower reach due to reduced numbers of PIT-tagged fish available, which may have affected the ability to identify ***the important factors".***

"In general, transportation provided benefits most years to Snake River hatchery spring/summer Chinook 1997-2004, ***however*** *benefits* ***varied*** *among hatcheries."*

"Migration year 2001 had very high ***but imprecise*** *TIRs, for both wild and hatchery steelhead."*

"Overall SARs for wild spring/summer Chinook fell short of the NPCC SAR objectives. Overall SARS of wild steelhead also fell short of NPCC SAR objectives although they exceeded those of wild Chinook. Based on these CSS SAR results relative to the NPPC SAR objectives, it appears that collecting juvenile fish at dams and transporting them downsteam in barges and trucks and releasing them downstream of Bonneville Dam did not compensate for the effect of the FCRPS on survival of wild Snake Basin spring/summer Chinook and steelhead migrating through the hydrosystem."

And finally the tacit assumption exists that differential post-Bonneville mortality between transported and in-river fish is "delayed mortality", i.e. an actual mortality event separated in time from its cause (once stated in the text specifically as "delayed mortality ***from*** *transport")*

We point out that : 1) whether or not the observed SAR in these years fell short of NPCC objectives provides no evidence one way or the other about compensating for the effects of the FCRPS; 2)the authors of the report have no knowledge of what the SAR would have been in these years if the FCRPS had not been in place; and 3) data now clearly provide the evidence that post-Bonneville mortality of transported fish is higher than for in-river migrants, but the reasons for this difference are still hypothetical.

<u>Response</u>: This NWFSC criticism is not well justified. The qualifying language (italicized by NWFSC) for the first three quotes accurately described our findings (identifying where transportation was beneficial, contrary to the NWFSC comment). For example, transportation did provide benefits most years to hatchery spring/summer Chinook, and benefits did vary among hatcheries. Also, TIR estimates for steelhead were imprecise in 2001. We have used neutral terms to describe results and implications of the CSS. Overall SARs from wild Snake River spring/summer Chinook and steelhead clearly have been less than NPCC objectives (minimum 2%, average 4%) across a wide range of ocean and migration conditions; whereas wild stream-type Chinook from downriver populations passing fewer dams have fared much better (see Figure 5.15). Post-Bonneville differential mortality between transported and in-river migrants is differential delayed mortality because it takes place after fish have transited the FCRPS. Moreover, our conclusion that transportation did not fully compensate for FCRPS effects is completely consistent with the NWFSC "Effects memo" (Williams et al. 2005) conclusion (p. xvi) that "transportation is not a panacea for negative effects of dams on fish stocks."

3. The authors repeatedly state that wild Chinook salmon do not meet the minimum 2% return rate goals of the region. Granted the CSS study uses only PIT-tagged fish, but in all cases where the comments on the 2% SAR goal are stated, no caveat exists that this represents data for PIT-tagged fish returns. The ISAB (2006) specifically indicated in comments on the 2005 CSS report that CSS participants needed to look into the potential disparity between PIT-tag returns and the unmarked population. Yet, in this report the ISAB comments are treated by a short discussion indicating that it was not clear how many actual wild spring-summer Chinook salmon passed Lower Granite Dam because some fish without ad-clips (ostensibly wild) were actually hatchery fish. Nonetheless, Copeland et al (2007) provided analyses of SARs for run-at-large nonad-clipped fish from the Snake River basin. In 3 of 5 years included in the CSS study (migration years 1998-2002, Figure 5.11), Copeland et al (2007) found that SARs exceeded 2% and more than 3.1% in 2 of them. They did not adjust for non-clipped hatchery fish in either the smolt or the adult life stages, so some bias in SARs may occur if differential survival existed between unmarked hatchery smolts and wild returns. Some unpublished analyses by NWFSC staff estimated the number of non-clipped hatchery smolts in the outmigration and used that to adjust adult returns to estimate numbers of wild fish (Figure 1). These analyses derived slightly different SARs than Copeland et al(2007) but they were similar.

<u>Response</u>: The introduction to Chapter 5 (p. 105) cites the ISAB (2006) issue that more attention should be given to whether PIT-tagged fish survive as well as untagged fish. Chapter 5 contains a section (p. 147) titled: “Do PIT-tag SARs represent SARs of the run-at-large?” with further discussion on p. 150-151. We agree with the ISAB (2006) conclusion that more attention should be given by CSS *and the Region as a whole* (emphasis added) to the discrepancy of SARs between PIT-tagged and untagged fish. However, the extremely tight reporting requirements did not allow for an examination of all the assumptions and data adjustments currently necessary to estimate SARs of the untagged component. Because the issue involves potential bias of both run-reconstruction and PIT-tag methodologies, resolution will require a collaborative effort among several technical groups in addition to the CSS project.

Contrary to the NWFSC comment that no caveat exists that PIT-tagged SARs may have a bias relative to the NPCC goal, the draft report explicitly stated (p. 147) “[t]he primary concern of negative bias from PIT-tag SARs would be in evaluating whether SARs are meeting NPCC biological objectives (2% minimum, 4% average).” Also, “[i]mplications of bias (if present) would be negligible for relative comparisons of the CSS PIT-tag SAR data, such as between Snake River migrants with different hydrosystem experiences, or between Snake River and downriver populations.” We also point to future monitoring and evaluation tasks to help resolve this issue in the future. We note that the 2 to 4 % goal itself was based on analyzes involving tagged fish that presumably experienced some handling mortality relative to the unmarked population.

Reviewer Comment: 4. Despite the ISAB recommendation to do so, this report does not include analyses of return rates of PIT-tagged and unmarked fish based on data in the CSS 2005 report (Berggren et al 2005). This seems most surprising given that the first four conclusions of this retrospective report laud the ability of the CSS group to PIT-tag over 2 million hatchery fish and analyze data from them. The absence of these analyses begs the question as to why and implies *the analyses may have wakened the reports statements about wild fish SARS. When NWFSC staff analyzed the CSS data we found that unmarked hatchery Chinook salmon returned at higher rates than PIT-tagged fish (Figure 2)which is similar to results from the analyses of wild Chinook Salmon and steelhead (Figure 1).*

Response: We addressed this issue in detail in Chapter 5. In addition, we also addressed this issue, in part, in Chapter 6, where we identify potential ways to address the question of PIT-tag detection and recovery at the hatchery weirs. Figure 2 of the NWFSC comments does not accurately represent hatchery-to-hatchery SARs of the PIT-tagged releases; the reviewers included a known negative SAR bias by including the bypassed group (C_1) as part of the PIT-tagged population, and by not weighting the C_0 and T_0 groups according to their actual proportions for the run at large. SARs of the C_1 category are substantially lower than those of C_0 (e.g., Figure 4.22), and the C_1 group is overrepresented in the NWFSC figure 2 analysis.

Reviewer Comment: 4. The reported SARs in this report are biased downward compared to standard SARs (eg Petrosky et al (2001)) because the authors base their SARs for Chinook salmon on adult returns only, not including jacks. This is important because the oft stated goal of reaching SArs of 2% is based on SArs that include jacks.

Response: The NPCC SAR goal was adapted from the 1998 PATH report (Marmorek et al. 1998). Comparison of model-generated median SARs and jeopardy probabilities (based on the NMFS interim standard for the 2000 BiOp) suggested median SARs must exceed 4% for the 48-year (interim) recovery standard, and 2% for the 100-year (interim) survival standard (Marmorek et al. 1998).

SARs may be calculated with or without jacks as recruits; there is no "standard" SAR. For most purposes, CSS has excluded jacks from the SAR calculations. However, a review of the 1998 PATH analysis indicates that jacks were included as recruits in the SARs, as noted by the reviewers. Therefore the CSS draft report contains a slight negative bias from this factor relative to the NPCC objective for spring/summer Chinook. Wild stream-type Chinook returns averaged only 4.2% jacks during the study period (Appendix D-39). Our initial comparison had the (quantitatively minor) inconsistency that we included jacks in the run-reconstruction estimates, which we have addressed. The run reconstruction SARs in the draft report inadvertently included jacks. This has been corrected to exclude jacks for consistency with the CSS SARs, and text has been modified. Methods and statistical assumptions for the CSS SARs are covered in Appendix B (and elsewhere) in the report. The inclusion of jacks in the SAR estimates would not change conclusions of the ten-year report regarding NPCC objectives because SARs missed the 2% NPCC minimum by such a wide margin.

Reviewer Comment: 5. The chapter deals extensively with within-season estimates of the following 4 quantities: water travel time (WTT), fish travel time (FTT), fish (cohort) survival (S), and "instantaneous mortality rate" (Z), which is derived as $S = \exp(-Z \cdot WTT)$ or equivalently, $\log(S)=-Z \cdot FTT$.

Response: This comment mischaracterizes our work on several levels. First, the comment reflects a fundamental misunderstanding of the differences between the dependent and independent variables that were analyzed. We analyzed three demographic rates as dependent variables: fish travel time, survival, and instantaneous mortality rates. We evaluated the degree of association between these dependent variables and seven independent variables: temperature, turbidity, flow, flow^{-1}, water travel time, average percent spill, and Julian day. Second, we defined the instantaneous mortality rate (Z) as $\hat{Z}_i = \frac{-\log_e(\hat{S}_i)}{F\hat{T}T_i}$, which is the maximum likelihood estimate for Z (Seber 1982:216). We did not equivocate WTT and FTT, as this commenter suggests, and this is a mischaracterization of our work. We found that FTT is a function of WTT, average percent spill, and Julian day, not just WTT as suggested by the commenter.

Reviewer Comment: 5. This formulation posits that a given cohort (as used here, weekly groups of fish arriving at Lower Granite Dam) has a particular instantaneous mortality rate and that direct survival through the hydropower system is directly related to fish travel time.

Response: First, the cohorts were defined as PIT-tagged fish detected and released into the Lower Granite Dam tailrace over a weekly time period, not weekly groups of fish *arriving* at Lower Granite Dam. Second, we estimated instantaneous mortality rates for weekly release cohorts through the equation defining the maximum likelihood estimate for Z, which is simply a transformation of the observed survival and median fish travel time rates. Third, we did not posit that weekly groups of fish have a particular instantaneous mortality rate upon arrival at Lower Granite Dam. Rather, that instantaneous mortality rates in each reach reflect the environmental or seasonal conditions experienced during migration through each reach. Predicted survival rates were then a function of the predicted instantaneous mortality rates and predicted fish travel times, both being functions of the environmental or seasonal conditions experienced during migration through each reach (termed "variable Z survival approach"). As an alternative analysis, we compared an approach where instantaneous mortality rates were at fixed levels within- and across-years, and that observed survival rates were primarily a function of changes in fish travel time (termed "constant Z survival approach"). We compare these two approaches, along with an approach that simply modeled survival rates as a function of environmental and seasonal conditions experienced during migration through the reach.

Reviewer Comment: 5. This formulation ignores that a substantial portion of the mortality occurs at the dams and is unrelated to fish travel time.

Response: The formulation used for instantaneous mortality rates accounts for differences in mortality rates that may occur during different periods during the migration. It reflects these differences as representing the arithmetic average mortality rates in cases where mortality rates may change over time (Keyfitz 1985:18-19).

Reviewer Comment: 5. As the authors note, FTT generally decreases within a season, and S (and log(S)) generally remains constant.

Response: While we found that FTT generally decreases over the migration season, there was substantial variation in survival rates over the migration season. There were examples of increasing survival trends, decreasing survival trends, and parabolic survival trends. Within-year survival rates could differ by up to 39 percentage points for both wild Chinook and steelhead, and by up to 32 percentage points for hatchery Chinook. We would not characterize survival rates as remaining constant within a season for either yearling Chinook or steelhead.

Reviewer Comment: 5. Thus, if two different groups of fish take a different amount of time to travel through a reach but their probability of surviving is the same, then per-day mortality of the two groups must be different.

<u>Response:</u> We would not disagree with this statement, as it follows from the inter-relationships between instantaneous mortality rates, survival rates, and time. However, this statement appears to imply that the instantaneous mortality rate is somehow a response variable, rather than the correct interpretation that it characterizes the average proportional mortality rate over time, essentially a transformation of observed survival rates and migration rates.

Reviewer Comment: 5. To conclude that decreasing FTT by managing the river to decrease WTT will result in increasing S (survival) requires the assumption that the quantity Z is an intrinsic characteristic of a group of fish; i.e., that the instantaneous mortality rate of the group is fixed at the time they leave Lower Granite Dam and that if we could only decrease their travel time to McNary Dam, then less mortality would occur.

<u>Response:</u> Again, this comment reflects some fundamental misunderstandings about our analyses. We did not assume that instantaneous mortality rates were fixed at the time they leave Lower Granite Dam. Rather, we assumed that instantaneous mortality rates reflected the environmental and/or seasonal conditions experienced *during migration* through the reach. Actions which may affect instantaneous mortality rates and/or actions which may affect fish travel times, both could affect resulting survival rates (under the variable Z survival approach). We also examined two other approaches (standard survival approach and constant Z survival approach) for predicting survival rates.

Reviewer Comment: 5. At least equally plausible and supported by observed data using the exact same relationship is a conclusion that management actions to decrease fish travel time would increase instantaneous mortality and that survival would remain the same.

<u>Response:</u> We have added a section to the discussion that examines this NWFSC hypothesis. To examine this hypothesis, we plotted the LGR-MCN instantaneous mortality rate estimates against observed median fish travel times for the early, mid, and late migration periods (Figure 2.23). We grouped the data by the early, mid, and late migration periods to account for potential seasonal differences in instantaneous mortality rates. An increase in instantaneous mortality rates as median fish travel times decrease would lend support to the NWFSC hypothesis. However, the data do not indicate that instantaneous mortality rates increase as median fish travel times decline (Figure 2.23). Based on the simple plots presented in Williams et al. (2005), which did not account for potential seasonal differences in instantaneous mortality, we understand how one might surmise that instantaneous mortality increases with decreasing fish travel times. However, we believe this is an incorrect interpretation of the data brought about by not accounting for the seasonal increases in instantaneous mortality that we frequently observed.

Reviewer Comment: 5. Therefore, the conclusion by the authors that decreasing FTT by half a day in the lower river would decrease steelhead mortality by 5.6% is highly questionable. Furthermore, the authors have incorrectly interpreted their result to derive this estimate. A Z of 0.112 does not imply a mortality of 11.2% per day. The correct interpretation is that the daily mortality is 1.0 – exp(-0.112), or 10.6%. Note that this discrepancy grows larger as FTT increases.

Response: The conclusion that decreasing FTT by half a day in the lower river would decrease steelhead mortality by 5.6% simply follows from the law of exponential population decline and the mean instantaneous mortality rates that were observed. Furthermore, for values of $Z \leq 0.1$, mortality rates and Z estimates are approximately equivalent (Ricker 1975). However, to clear up any confusion on the trivial differences between the two, we have provided both daily percent mortality estimates and Z estimates (Tables 2.1, 2.2).

Reviewer Comment: 5. When the authors relate Z to a variety of factors, an additional problem is encountered. WTT and FTT are correlated with each other and relatively stable within seasons, and as stated above, S (and log(S)) has repeatedly remained relatively constant within seasons, especially for spring-summer Chinook salmon. The final quantity (Z) is derived by dividing the relatively constant quantity log(S) by the relatively variable FTT. It is no surprise, then, that Z and WTT are correlated. In fact, this is inevitable because of the relationships described above and is a classic example of a "spurious correlation."

Response: First, consistent with Williams et al. (2005), we examined the relationship between instantaneous (daily) mortality rates and water travel time (along with five other independent variables). Criticisms levied the NWFSC for our examination of the relationship between instantaneous mortality rates and WTT, when the NWFSC has conducted similar analyses (Williams et al. 2005), are hypocritical. Second, with the correlation between WTT and FTT, one must remember which is a considered a response variable (FTT) and which is considered an independent variable (WTT). FTT cannot influence WTT, whereas WTT may or may not influence FTT. We found that several other independent variables (average percent spill and Julian day), not just WTT, influenced FTT. Third, we observed some fairly dramatic increasing, decreasing, and parabolic seasonal trends in within-season estimates of survival. Within-season survival rates could differ by up to 39 percentage points for both wild Chinook and steelhead, and by up to 32 percentage points for hatchery Chinook. We would not characterize survival rates as remaining constant within a season for either yearling Chinook or steelhead. The instantaneous mortality rates (*Z*) largely reflected these changes in survival rates, with most of the variation in instantaneous mortality rates associated with variation in survival (49% for Chinook and 58% for steelhead), followed by Julian day (35-36% for Chinook and steelhead) (Table 2.11).

Reviewer Comment 6. Comments regarding attention on wild vs. hatchery fish, use of C_0 vs. C_1 fish, and evidence indicates only that there is no benefit to transporting wild Chinook, not that it is harmful.

Response: In the report, we did look at temporal (within-season) variation in SARs in Chapter 4, using C_1 fish as surrogates. Further, annual estimates can be useful in comparing seasonal transportation modification strategies, under an adaptive management regime (i.e. change strategy, monitor how annual SARs, TIRs, *D*s change from the "baseline").

In a sense, CSS C_0 fish are not represented by reach survival rate estimates of tagged fish, due to different disposition at dams. However, the CJS model requires downstream recaptures (detections) in order to estimate detection probability and survival rates. Therefore, the assumption that detection history doesn't affect significantly affect short reach survival rates is necessary for survival rate estimation. If violation of this assumption is influential, all reach survival estimates (including NOAA's) are affected.

Chapter 3 provides extensive results for SARs, TIRs, and *D* estimates for hatchery Chinook and steelhead. Absolute values and trends in these quantities are compared between wild and hatchery fish. Chapter 4 suggests that transportation, as currently implemented, is detrimental to wild Chinook, since a majority of the TIR distribution at each project falls below one.

6. Chapter 3 NOAA Comment (Part A): The chapter focuses mostly on wild Chinook salmon, and therefore does a poor job of comparing the results of analyses among wild and hatchery Chinook salmon, and wild and hatchery steelhead. Without these comparisons, managers have little ability to determine the best strategies that will lead to the optimum return for the different species and type (wild or hatchery).

Response: Based on all comments from all reviewers of Chapter 3, a major rewrite of the results and discussion section of this Chapter has rectified those concerns.

Chapter 3 NOAA Comment (Part B): Another shortcoming of the analysis derives from the authors' insistence on only using C_0 fish as "true controls.' They argue that because these fish are not seen at transport dams, no temporal analyses are possible. Thus, the analyses presented in this chapter will provide little guidance on the important management questions for each transport dam related to when to begin transportation within a season, and when and how much spill should occur. The emphasis on "true controls" in the CSS study seems misplaced. A better foundation for analyses would use data similar to what is presented in Table 5.16. Here, data comparing C_0 to C_1 fish (for fish observed at Bonneville Dam) indicate that in the preponderance of comparison, C_1 fish have equivalent SARs of the C_0 fish (point estimates in most years for bi-weekly comparisons are higher). These are the fish that make it successfully to Bonneville Dam from the different categories. Thus, it appears that use of C_1 fish would provide some useful insight into temporal changes in return rates of transported and non-transported fish. Analyses along this line would significantly improve this chapter.

Response: The wording "true controls" for C_0 fish has been removed from the text. The C_0 group is the closest representation of the untagged run-at-large fish that are not transported from the three Snake River collector dams during the years analyzed in this report. With the exception of 1997 when a management operation of bypassing most untagged steelhead at LGS and LMN throughout the season was attempted, the other years analyzed in this report (1994-1996 and 1998-2004) where periods when the management operation was to transport all collected untagged run-at-large fish. In the estimation of TIR, we are evaluating the operational condition whereby untagged run-at-large fish are transported if collected relative to those untagged run-at-large fish not collected. Therefore using the PIT-tagged groups that closest reflect those two groups are proper choices for the TIR estimation. If the question had been what to do with the collected fish, then using SAR(T_0) and SAR(C_1) in the TIR estimation would have been proper. The question of temporal changes in SARs was not covered in Chapter 3, but is covered for wild Chinook and wild steelhead in Chapter 4 using dam-specific estimates of transported and bypassed PIT-tagged fish. Whether one uses C_0 or C_1 fish in an particular evaluation must be determined by the question at hand though, and not by whether post-BON SAR estimates for groups C_0 and C_1 are similar, as inferred by NOAA in the latter part of their comment regarding data from Table 5.16. PIT-tagged fish in Table 5.16 are fish that survived to the lower river, whereas the PIT-tagged fish used in the CSS estimations of TIR and D are based on estimated numbers of T_0 and C_0 fish beginning their passage through the hydrosystem.

Chapter 3 NOAA Comment (Part C): Additionally, nearly all the analyses discussed presume that survival estimates for non-transported fish (the "true controls") are the same as those of the marked population used to make juvenile survival estimates. However, using the CSS argument, the PIT-tagged fish returned to the river do not represent "true controls" and do not measure the survival of fish not detected at transport dams because they are based on the combined population of detected and non-detected fish. A disconnect thus occurs. Since non-detected fish mostly pass through spill, one might reasonably assume they have a higher survival than the combined population.

Response: In the estimation of in-river reach survival rates between the dams with detectors, all users (including NOAA) of PIT-tagged data in the Columbia River basin have had to rely on the assumption that prior detection history is not influencing subsequent detection probabilities and reach survival rates when using the CJS model to estimate those reach survival rates. NOAA is trying to paint the picture that since we do not use C_1 fish as "true controls," then we should not be using C_1 fish in the estimation of reach survival rates. As stated earlier, the term "true controls" is misleading since the proper in-river group to use in any comparison will be determined by the question being answered. There is no such thing as a "true control" for every analysis. That said, NOAA raises a legitimate concern that has ramification for all users of PIT-tag data (including NOAA themselves) within the Columbia River basin for reach survival estimation. It is generally accepted based on years of COE funded evaluations of survival through spillways, bypasses, and turbines, that the spillway route gives a higher survival than bypass route. Therefore, when using the CJS model to estimate a common parameter of survival for a particular reach, all researchers (including NOAA) need to realize that each inter-dam reach survival rate estimate encompasses the unmeasured components of reservoir survival rate times weighted average of route-specific survival rate across the routes of spillway, bypass, and turbine, where the weights are the proportion of the population of PIT-tagged fish utilizing each of these three routes through a project. But in using the CJS model, we, NOAA, and others accept the assumption that all PIT-tagged fish used in estimating a particular reach survival rate are independently and identically distributed about a common reach survival rate for that particular reach. If a "disconnect" exists as stated by NOAA, then they too are part of that disconnect.

Chapter 3 NOAA Comment (Part D): Finally, even the data presented I the CSS study, when considered on an annual basis, do not indicate that transportation harms wild Chinook salmon;

just that it provides no benefit. The annual data for hatchery Chinook and steelhead all show a substantial benefit that would potentially translate into thousand of additional adult returns if spilling or collecting and transporting fish were optimized for all species at each dam. Caution on potential benefits for hatchery Chinook is warranted, however, as the CSS associated hatcheries and numbers of PIT-tagged fish released from each do not mirror the total hatchery production released in the basin.

Response: We report that the SAR data from 1994 to 2004 does not appear to show a benefit of transportation except in drought years such as 2001. The CSS did show and acknowledge transportation benefits to four of the five hatcheries used in the CSS (Rapid River, McCall, Imnaha, and Catherine Ck, but not Dworshak), and for wild and hatchery steelhead. However, delayed differential mortality of transported fish compared to the in-river migrants dampens the potential that may be achieved by transportation alone as a management tool aimed at recovering listed fish. We do not claim that the five hatcheries above LGR used in the CSS reflect all of hatchery production. Since we see differences in response to transportation among the five hatcheries used in the CSS, which currently account for approximately half of production of spring/summer Chinook from hatcheries above LGR, it is likely differences in response to transportation will also occur across those remaining hatcheries.

Reviewer Comment 7. The graphs in Chapter 4 always indicate the 2% SAR line when the majority of estimates fall below the line, but often do not include the 2% SAR line when the majority of estimates fall above it.

Response: The 2-6% desired range of SARs adopted by the NPCC was originally developed for Chinook, rather than steelhead. At the time of some of the analyses, the author of Chapter 4 was uncertain whether the target had been adopted for steelhead as well, so these weren't included in some of the steelhead figures (though the 2-6% target range was included in the aggregate steelhead SAR figure). In the rush to meet the deadline for posting the draft report, standardization of all figures was not a priority. In Chapter 4 of the revised report, the 2-6% range is indicated on all SAR figures, with the exception of the within –season figures (to avoid clutter).

Reviewer Comment: 8. The continued emphasis by CSS to compare upstream/downstream population productivity appears misplaced and has limited utility for estimating overall hydropower system impacts. We concur with the conclusion of the ISAB latent Mortality Report (2007) which stated "The ISAB concludes that the hydrosystem causes some fish to experience latent mortality, but strongly advises against continuing to try to measure absolute latent mortality. Latent mortality relative to a damless reference is not measurable. Instead, the focus should be on the total mortality of the in-river migrants and transported fish, which is the critical issue for recovery of listed salmonids. Efforts would be better expended on estimation of processes such as in-river versus transport mortality that can be measured directly."

In addition the ISABs comments and flaws of the upstream/downstream approach that have been identified previously (Zabel and Williams 2000; Williams et al 2005), we provide two additional comments;

- *Weak scientific methodology. The standard scientific method operates by stating a null and alternative hypotheses and considering all available information in an effort to reject the null hypotheses. Science does not wok by laying out a hypothesis then saying it is correct unless positive proof exists to show that it is wrong. Yet, this is what has occurred here.*
- *Ignores data from other systems. Data on natural sockeye salmon populations in Bristol Bay have shown similar trends in overall productivity as have the upstream/downstream comparisons used by CSS. Overall productivity of the Bristol Bay populations increased and decreased over a period of decades, concomitant with major changes in ocean conditions. However, some of these eight closely related populations demonstrated strikingly divergent temporal patterns (Hilborn et al, 2003; Peterman et al. 2003). Yet the analyses comparing Snake River and John Day River Chinook salmon populations assume that changes in temporal patterns do not exist. The Bristol Bay data suggest a lack of foundation for this assumption.*

Response: One major objective of the CSS study was to "begin a time series of SARs for use in hypothesis testing and in the regional long-term monitoring and evaluation program". The intent was not to limit analyses to one particular statistical model. CSS did lay out several null hypotheses and the study was designed to address these, e.g., through estimating number of marked fish in each group to achieve target confidence levels that *TIR* was > 1. The hypotheses were framed as in the 1996-98 CSS status report (CSS 2000): "Test if the annual ratio of transport survival rate to in-river survival rate (measured at Lower Granite Dam) is greater than 1.5 with sufficient power to provide a high probability that the ratio is greater than 1.0." The "standard scientific method" with null and alternative hypotheses is hardly the only way that applied science is conducted. CSS has tested particular hypotheses under the null/alternative hypothesis formulation (e.g. see below), but has also performed parameter estimation, especially confidence interval estimation, and model selection. There is much applied science done outside of the traditional null/alternative hypothesis formulation in other ways, too; e.g. model selection, estimation of Bayesian credibility intervals, formal decision analysis, etc.

We are confused by the reviewers' characterization of the CSS analysis in this comment. Contrary to NWFSC comment, we clearly stated that the purpose (p. 106) of the upriver/downriver SAR comparison was to determine if the difference in mortality estimated from spawner-recruit (SR) analyses was also apparent in the SARs (i.e., H_0: differential mortality from SARs equals differential mortality from SR). Contrasts of the point estimates and 90% CI from the two types of data (p. 131-133) indicated SAR-based estimates of differential mortality agreed well with published SR-based estimates of differential mortality. We characterized the upriver-downriver comparison as a "natural experiment", which therefore has some design limitations (p. 150). Further, we investigated and tested hypotheses regarding possible non-hydrosystem causes (including alternative hypotheses previously suggested by NWFSC) of differential mortality between upriver and downriver wild stream-type Chinook (p. 136-143).

Based on 5 years of PIT-tag SAR comparisons between wild Snake River and John Day smolts, we have seen a consistent pattern of differential mortality across poor and favorable ocean conditions. Combined with estimates of in-river survival and relative survival of transported smolts, this is one line of (indirect) evidence that the magnitude of delayed hydrosystem mortality is large (e.g., Peters and Marmorek 2001; Schaller and Petrosky 2007). However, actual estimation of delayed or latent mortality (of in-river migrants) was not an objective of CSS, and we did not attempt to estimate it in the CSS draft 10-year report, contrary to the reviewers' comments.

In addition to the upriver-downriver comparison, we investigated the influence of ocean/climatic and migration conditions on SARs of wild spring/summer Chinook in Chapter 5. Water travel time (WTT), a measure of water velocity through a fixed reach, was influential in all top multiple regression models (p. 128-131); May or September PDO were also typically incorporated in top models. The coefficients for WTT vs. ln(SAR) were consistent across models, ranging from -0.053 to -0.076. That is, for each day increase in WTT, the SAR would be expected to decrease 5% - 8%, or 65%-78% for a 20 day increase in WTT. This result is generally consistent with the differential mortality estimated from upriver-downriver comparison of wild Chinook, and was an important independent estimate that did not rely on the use of downriver reference populations.

Contrary to the NWFSC reviewers' comments, we have previously examined data from other systems, including the Bristol Bay dataset, which the reviewers claim invalidates comparing performance of different populations from the same region. We don't agree. Pyper et al. (2005) incorporated this stock group in their analysis, and found correlations in survival rate patterns up to 500 km from the ocean point of entry (upriver and downriver stocks in our analysis have the same point of ocean entry). Schaller and Petrosky (2007) found that variation of survival rates (SR residuals) of Snake River stream-type Chinook were more variable than those from than most other stock groups used in Pyper et al. (2005). Specifically, Snake River populations showed significantly greater variability in survival rate indices than the Bristol Bay group ($F=3.42$, $p<0.0001$). We plotted the mean and range of the SR residuals for the Bristol Bay sockeye stock group in Figure 1 below (data from R. Peterman and B. Pyper, personal communication). Even within the diverse complex of Bristol Bay sockeye salmon, there are discernable annual survival rate patterns (Figure 1); correlations between sockeye stocks within the Bristol Bay stock group ranged from 0.23 to 0.75 (geometric mean 0.44).

Further, the reviewers' reference to Hilborn et al. (2003) failed to identify that many of the differences within the Bristol Bay sockeye salmon complex were attributed to varying challenges imposed by the different freshwater spawning and rearing environments (e.g., lakes, rivers, and streams). The upriver and downriver stream-type Chinook compared in CSS (and previous SR contrasts) have more similar freshwater life-history characteristics than the Bristol Bay sockeye. The situation in the Columbia River stream-type Chinook SR analyses is that these papers (Schaller et al. 1999; Deriso et al. 2001; Schaller and Petrosky 2007) explicitly compared populations from stream spawning and rearing fish, where we specifically accounted for differences in freshwater carrying capacity and productivity in the SR analysis (given that we have stream specific spawner, age structure, and recruit information). In any case, Bristol Bay sockeye data do not support the implied criticism that variability in ocean survival among groups could create the false impression of systematic differences between groups of sockeye.

The present CSS comparison extends the SR analyses (and provides an independent estimate of differential mortality that does rely on assumptions for a particular recruit/spawner function) by estimating differential mortality based on PIT-tag SARs, and also by examining specific life-history characteristics which might support alternative hypotheses regarding causes of differential mortality. Our approach is consistent with the recommendations of Hilborn et al. (2003) in that analysis should be applied on a scale where one can estimate stream-specific recruit/spawner ratios and survival rates.

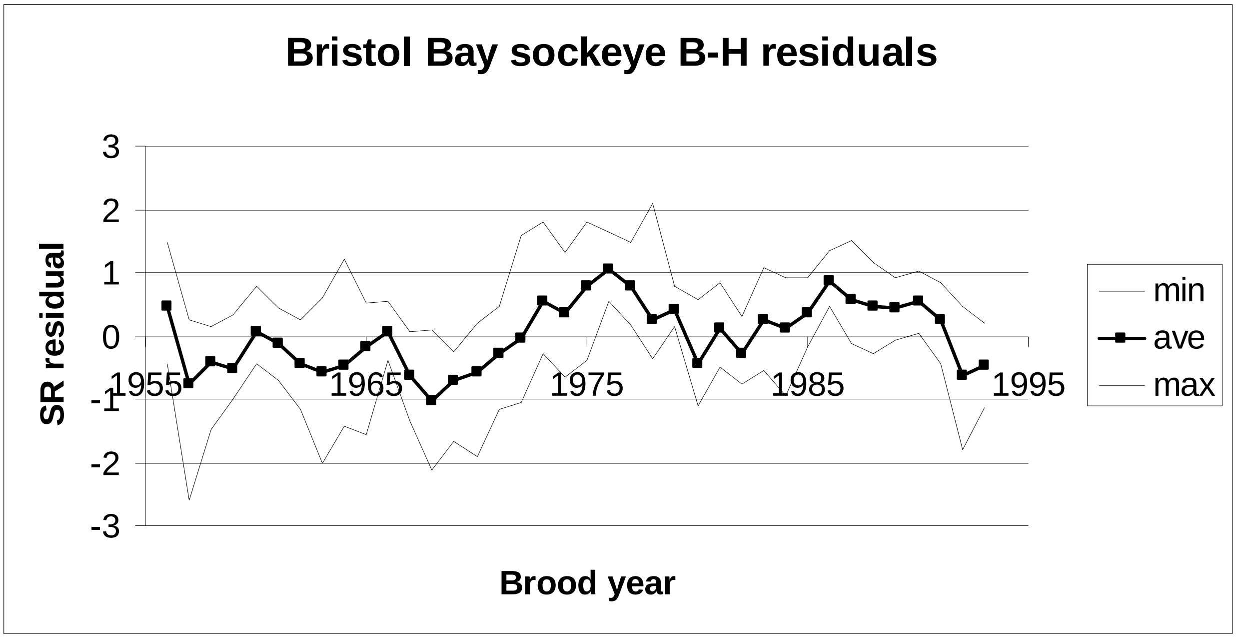


Figure 1. Minimum, mean and maximum annual spawner-recruit residuals for Bristol Bay populations from Pyper et al. 2005 (R. Peterman and B. Pyper, pers. comm.).

Reviewer Comment 9: No clear direction exists to argue for continuing the large releases of hatchery fish for the purposes of 'comparative' survival. This is based on: 1) It does not appear that hatchery Chinook salmon provide any useful information related to wild Chinook salmon other than when SARs for hatchery Chinook salmon go way up or way down, proportionately, so do SARs for wild Chinook salmon. This could be determined from a much smaller number of PIT-tagged fish or from adult returns by comparing the clipped to unclipped population. 2) The CSS results indicate that on an annual basis, transportation would benefit hatchery Chinook salmon but not wild Chinook salmon. Since the distribution of hatchery Chinook salmon past lower Granite Dam is much more compressed than that of wild Chinook salmon, it is not clear that even analyses on a temporal basis with hatchery Chinook salmon would provide information on how best to operate the system for wild Chinook salmon. 3) Hatchery Chinook salmon have a wide range in return rates. McCall fish do particularly well, and have a different distribution than Dworshak fish. Which hatchery fish then represent wild fish?

<ins>Response</ins>: Hatchery Chinook salmon and wild Chinook salmon responded nearly identically to environmental and/or seasonal conditions in terms of their fish travel time, instantaneous mortality rates, and survival rates in the LGR-MCN reach. Thus, hatchery Chinook salmon

provide valuable information on the response of wild Chinook salmon to conditions experienced in the hydrosystem.

Differential mortality between upriver and downriver stream-type Chinook populations has been estimated for wild populations from both spawner-recruit (Schaller et al. 1999; Deriso et al. 2001; Schaller and Petrosky 2007) and PIT-tag SAR (CSS study) data sources. The CSS also investigated whether a similar level of differential mortality was present between PIT-tag SARs for five upriver and one downriver hatchery Chinook populations. Because biological characteristics of a population could differentially influence survival to adult return (see above), we also summarized hatchery pre-smolt FL at the time of tagging, and hatchery smolt arrival timing distributions entering the hydrosystem (LGR or BON) and arriving at the estuary (BON).

Upriver and downriver hatchery spring/summer Chinook SARs did not show the same level of differential mortality as was apparent from the wild populations. Survival of hatchery fish is subject to additional fitness and rearing factors that may not affect wild populations. CSS currently has the ability to compare SARs from a single downriver hatchery (Carson NFH) with those from five Snake River hatcheries. Additional candidate populations relevant to these SAR comparisons from downriver hatcheries of the Interior Columbia include Klickitat, Warm Springs, and Round Butte (depending on fish health constraints). Future monitoring should also consider incorporating PIT-tag SARs from the upper Columbia region to expand these regional comparisons.

Although Snake River hatchery Chinook exhibited a generally more positive response to transportation and relatively lower levels of differential mortality than wild populations, annual SARs of wild and hatchery Snake River Chinook were highly correlated. In view of this high correlation, continuing the CSS time series of hatchery SARs will be important to augment wild Chinook SAR information following future years of low escapements, in addition to providing valuable management information for the specific hatcheries. One advantage of the CSS study is that tagging takes place at the hatcheries and in the tributaries for wild populations. This approach allows for detecting different responses to management actions for different components of the wild and hatchery aggregate groups, unlike approaches that only tag at the upper most dam. Finally, it is of interest to the region of how the specific hatchery groups respond to the hydrosystem management actions. The reviewers suggest a much smaller number of PIT-tagged hatchery fish could be used. We believe that the sample sizes should be periodically reviewed based on updated survival estimates, and regional monitoring and evaluation needs.

UNITED STATES DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
NATIONAL MARINE FISHERIES SERVICE
1201 NE Lloyd Boulevard, Suite 1100
PORTLAND, OREGON 97232-1274

F/NWR5

July 2, 2007

Comparative Survival Study Oversight Committee
c/o Fish Passage Center
1827 NE 44th Street, Suite 240
Portland, OR 97213

Re: Comments on 10-year draft CSS report

Dear Committee Members:

Thank you for the opportunity to provide comments on the draft 10-year Comparative Survival Study (CSS) report. Our Northwest Regional Office asked the Northwest Fisheries Science Center (NWFSC) to review the CSS draft report with the following issues in mind:

- Does the CSS report come to different conclusions on the effects of flow, spill, and temperature on survival of yearling migrants than NWFSC's 2005 *Effects* memo?
- Does the CSS report come to a different conclusion on the benefits of transport than the NWFSC's 2005 *Effects* memo?
- Did the CSS report draw conclusions on issues on latent mortality that the recent ISAB report considered unsupportable?
- Was the CSS report too narrow in its analysis of factors affecting SARs?
- Were there conclusions drawn by the CSS report that the NWFSC would take issue with because the methods used to study the issue were not valid?

Enclosed is the response provided. Thank you again for the opportunity to comment on the draft report.

Sincerely,

Bruce K. Suzumoto
Assistant Regional Administrator
Hydropower Division

Enclosure

cc: Tony Grover, Northwest Power and Conservation Council

UNITED STATES DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
NATIONAL MARINE FISHERIES SERVICE
Northwest Fisheries Science Center
2725 Montlake Boulevard East
Seattle, WA 98112-2097

June 29, 2007

MEMORANDUM TO: D. Robert Lohn
Regional Administrator, NW Region

FROM: Usha Varanasi, Ph.D.
Science and Research Director, NWFSC

SUBJECT: Comment on the DRAFT Comparative Survival Study (CSS) 2007 Ten-year Retrospective Analysis Report

At the request of Paul Wagner and Ritchie Graves, we reviewed the DRAFT "Comparative Survival Study (CSS) of PIT-tagged Spring/Summer Chinook and Steelhead in the Columbia River Basin Ten-year Retrospective Analyses Report." The report is extraordinarily long (377 pages); too long to read, digest, and provide finely detailed commentary in the review time available. The following paragraphs summarize our major concerns with the report. Please call John Williams (206.860.3277) if you have any questions regarding these comments.

1. Most strikingly, despite its title and the fact that the CSS study group has PIT tagged hundreds of thousands of juvenile Chinook salmon and steelhead, the CSS retrospective report does not contain a holistic analysis of this 10-year effort or an integration of the results across all species that considers different migration conditions.

2. The data presented, and the discussion and conclusion sections all seem focused through the lens of specific positions favored by the authors: hydropower system-related latent mortality is large in magnitude, transportation is not beneficial, management actions directed at the hydropower system have generally failed, and consequently SARs have been low in recent years and drastic actions are needed to recover the wild Chinook salmon populations, as PIT-tagged wild fish fail to meet a minimum 2% SAR. Results that do not support desired positions are usually discounted by carefully placed language. For example, from the conclusions in Chapter 8 (*all italics are ours*):

"Variation in [survival] in the MCN-BON reach was explained by temperature and Julian day. However, there was substantial uncertainty in the lower reach due to reduced numbers of PIT-tagged fish available, which may have affected the ability to identify *the important factors.*"

"In general, transportation provided benefits most years to Snake River hatchery spring/summer Chinook 1997-2004, *however* benefits *varied* among hatcheries."

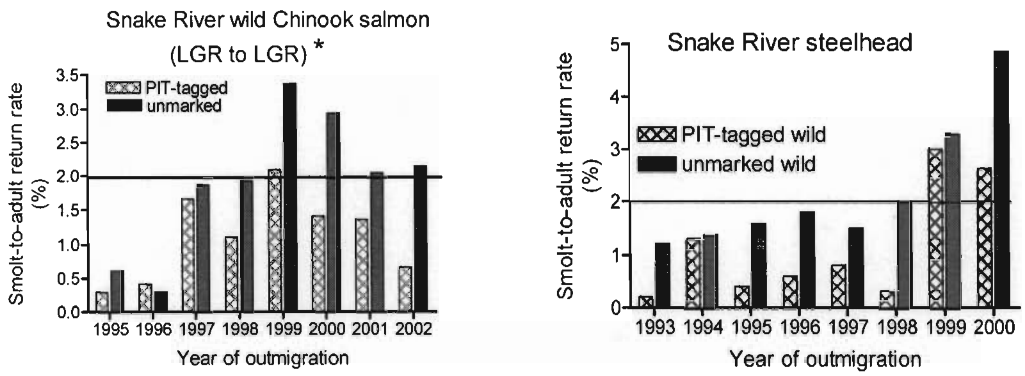
"Migration year 2001 had very high, *but imprecise* TIRs, for both wild and hatchery steelhead."

"Overall SARs for wild spring/summer Chinook fell short of the NPCC SAR objectives. Overall SARs of wild steelhead also fell short of NPCC SAR objectives, although they exceeded those of wild Chinook. Based on these CSS SAR results relative to NPPC SAR objectives, it appears that collecting juvenile fish at dams and transporting them downstream in barges and trucks and releasing them downstream of Bonneville Dam did not compensate for the effect of the FCRPS on survival of wild Snake Basin spring/summer Chinook and steelhead migrating through the hydrosystem."

And finally, the tacit assumption exists that differential post-Bonneville mortality between transported and in-river fish is "delayed mortality", i.e., an actual mortality event separated in time from its cause (once stated in the text specifically as "delayed mortality *from* transport".)

We point out that: 1) whether or not the observed SAR in these years fell short of NPCC objectives provides no evidence one way or the other about compensating for effects of the FCRPS; 2) the authors of the report have no knowledge of what the SAR would have been in these years if the FCRPS had not been in place; and 3) data now clearly provide the evidence that post-Bonneville mortality of transported fish is higher than for in-river migrants, but the reasons for this difference are still hypothetical.

3. The authors *repeatedly* state that wild Chinook salmon do not meet the minimum 2% return rate goals of the region. Granted the CSS study uses only PIT-tagged fish, but in all cases where the comments on the 2% SAR goal are stated, no caveat exists that this represents data from PIT-tagged fish returns. The ISAB (2006) specifically indicated in comments on the 2005 CSS report that CSS participants needed to look into the potential disparity between PIT-tag returns and the unmarked population. Yet, in this report the ISAB comments are treated by a short discussion indicating that it was not clear how many actual wild spring-summer Chinook salmon passed Lower Granite Dam because some fish without ad-clips (ostensibly wild) were actually hatchery fish. Nonetheless, Copeland et al. (2007) provided analyses of SARs for run-at-large nonad-clipped fish from the Snake River basin. In 3 of the 5 years included in the CSS study (migration years 1998-2002, Fig. 5.11), Copeland et al. (2007) found that SARs exceed 2%, and more than 3.1% in 2 of them. They did not adjust for non-clipped hatchery fish in either the smolt or adult life stages, so some bias in SARs may occur if a differential survival existed between unmarked hatchery smolts and wild returns. Some unpublished analyses by NWFSC staff estimated the number of non-clipped hatchery smolts in the outmigration and used that to adjust adult hatchery returns to estimate numbers of wild fish (Figure 1). These analyses derived slightly different SARs than Copeland et al. (2007), but they were similar.

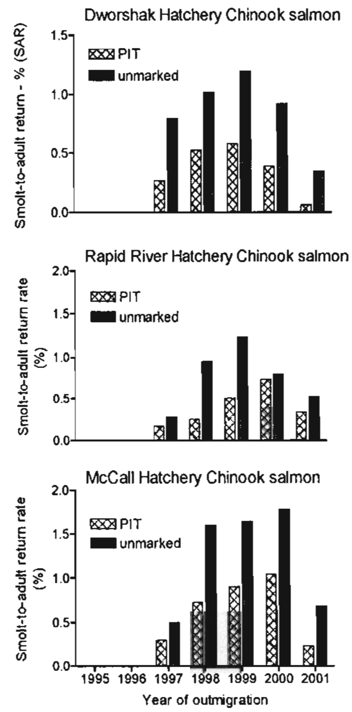


* Not adjusted for harvest

Figure 1. Comparison of SARS for PIT-tagged and run-at-large Snake River wild spring-summer Chinook salmon and steelhead.

Clearly, these analyses indicate higher SARs for unmarked wild fish compared to PIT-tagged wild spring-summer Chinook salmon. The geometric return rate of PIT-tagged fish was only 60% that of wild fish comparing CSS results to those of Copeland et al. (2007). These, however, are not the only data that exist indicating differences in return rates between PIT-tagged and unmarked fish. Petrosky (unpublished data and an author of this report) estimated SARs for the unmarked population of Snake River wild steelhead and submitted this to the Interior Columbia River Technical Recovery Team. The NWFSC's analysis that compared SARs of these fish to an estimated SAR for the PIT-tagged population (no data for the unmarked population exists beyond the 2000 outmigration) also showed in all years PIT-tagged fish returned at rates less than the unmarked population (Figure 1).

Despite the ISAB recommendation to do so, this report does not include analyses of return rates of PIT tagged and unmarked fish based on data in the CSS 2005 report (Berggren et al. 2005). This seems most surprising given that the first four conclusions of this retrospective report laud the ability of the CSS group to PIT tag over 2 million hatchery fish and analyze data from them. The absence of these analyses begs the questions as to why, and implies the analyses may have weakened the report's statements about wild fish SARs. When NWFSC staff analyzed the CSS data we found that unmarked hatchery Chinook salmon returned at higher rates than PIT-tagged fish (Figure 2), which is similar to results from the analyses of wild Chinook salmon and steelhead (Figure 1).



* Includes adjustment for harvest upstream of LGR

Figure 2. Hatchery to hatchery SARs (no adjustment for differences in downstream migration history for PIT and unmarked fish). Data after Berggren et al. (2005).

4. The reported SARs in this report are biased downward compared to standard SARs (e.g., Petrosky et al. (2001)) because the authors base their SARs for Chinook salmon on adult returns only, not including jacks. This is important because the oft stated goal of reaching SARs of 2% is based on SARs that include jacks.

5. Chapter 2. The chapter deals extensively with within-season estimates of the following 4 quantities: water travel time (WTT), fish travel time (FTT), fish (cohort) survival (S), and "instantaneous mortality rate" (Z), which is derived as $S = \exp(-Z \cdot WTT)$ or equivalently, $\log(S)=-Z \cdot FTT$. This formulation posits that a given cohort (as used here, weekly groups of fish arriving at Lower Granite Dam) has a particular instantaneous mortality rate and that direct survival through the hydropower system is directly related to fish travel time. This formulation ignores that a substantial portion of mortality occurs at the dams and is unrelated to fish travel time. As the authors note, FTT generally decreases within a season, and S (and log(S)) generally remains constant. Thus, if two different groups of fish take a different amount of time to travel through a reach but their probability of surviving is the same, then the per-day mortality of the two groups must be different. To conclude that decreasing FTT by managing the river to

decrease WTT will result in increasing S (survival) requires the assumption that the quantity Z is an intrinsic characteristic of a group of fish; i.e., that the instantaneous mortality rate of the group is fixed at the time they leave Lower Granite Dam and that if we could only decrease their travel time to McNary Dam, then less mortality would occur.

At least equally plausible and supported by observed data using the exact same relationship is a conclusion that management actions to decrease fish travel time would increase instantaneous mortality and that survival would remain the same. With respect to Chinook salmon, a more parsimonious explanation for the observed results is that most of the estimated mortality in the LGR-MCN reach occurs during passage at dams, independent of flow, WTT, and FTT, and very little occurs in the reservoirs themselves. Available survival data for dam passage from both PIT-tag and radio-tag studies for Chinook salmon lend more support to dams as the area where changes in survival occur. Therefore, the conclusion by the authors that decreasing FTT by half a day in the lower river would decrease steelhead mortality by 5.6% is highly questionable. Furthermore, the authors have incorrectly interpreted their result to derive this estimate. A Z of 0.112 does not imply a mortality of 11.2% per day. The correct interpretation is that the daily mortality is 1.0 – exp(-0.112), or 10.6%. Note that this discrepancy grows larger as FTT increases.

When the authors relate Z to a variety of factors, an additional problem is encountered. WTT and FTT are correlated with each other and relatively variable within seasons, and as stated above, S (and log(S)) has repeatedly remained relatively constant within seasons, especially for spring-summer Chinook salmon. The final quantity (Z) is derived by dividing the relatively constant quantity log(S) by the relatively variable FTT. It is no surprise, then, that Z and WTT are correlated. In fact, this is inevitable because of the relationships described above and is a classic example of a "spurious correlation."

6. Chapter 3. The chapter focuses mostly on wild Chinook salmon, and therefore does a poor job of comparing the results of analyses among wild and hatchery Chinook salmon, and wild and hatchery steelhead. Without these comparisons, managers have little ability to determine the best strategies that will lead to the optimum return for the different species and types (wild or hatchery).

Another shortcoming of the analyses derives from the authors' insistence on only using C_0 fish as "true controls". They argue that because these fish are not seen at transport dams, no temporal analyses are possible. Thus, the analyses presented in this chapter will provide little guidance on the important management questions for each transport dam related to when to begin transportation within a season, and when and how much spill should occur. The emphasis on 'true controls' in the CSS study seems misplaced. A better foundation for analyses would use data similar to what is presented in Table 5.16. Here, data comparing C_0 to C_1 fish (for fish observed at Bonneville Dam) indicate that in the preponderance of comparisons, C_1 fish have equivalent SARs of the C_0 fish (point estimates in most years for most bi-weekly comparisons are higher). These are the fish that make it successfully to Bonneville Dam from the different categories. Thus, it appears that use of C_1 fish would provide some useful insight into temporal changes in return rates of transported and non-transported fish. Analyses along this line would significantly improve this chapter.

Additionally, nearly all the analyses discussed presume that survival estimates for non-transported fish (the "true controls") are the same as those of the marked population used to make juvenile survival estimates. However, using the CSS argument, the PIT-tagged fish returned to the river do not represent "true controls" and do not measure the survival of fish not detected at transport dams because they are based on the combined population of detected and non-detected fish. A disconnect thus occurs. Since non-detected fish mostly pass through spill, one might reasonably assume they have a higher survival than the combined population.

Finally, even the data presented in the CSS study, when considered on an annual basis, do not indicate that transportation harms wild Chinook salmon; just that it provides no benefit. The annual data for hatchery Chinook salmon and steelhead all show a substantial benefit that would potentially translate into thousands of additional adult returns if spilling or collecting and transporting fish were optimized for all species at each dam. Caution on potential benefits for hatchery Chinook salmon is warranted, however, as the CSS associated hatcheries and numbers of PIT-tagged fish released from each do not mirror the total hatchery production released in the basin.

7. The graphs in Chapter 4 always indicate the 2% SAR line when the majority of estimates fall below the line, but often do not include the 2% SAR line when the majority of estimates fall above it.

8. The continued emphasis by CSS to compare upstream/downstream population productivity appears misplaced and has limited utility for estimating overall hydropower system impacts. We concur with the conclusion of the ISAB Latent Mortality Report (2007) which stated "The ISAB concludes that the hydrosystem causes some fish to experience latent mortality, but strongly advises against continuing to try to measure absolute latent mortality. Latent mortality relative to a damless reference is not measurable. Instead, the focus should be on the total mortality of the in-river migrants and transported fish, which is the critical issue for recovery of listed salmonids. Efforts would be better expended on estimation of processes, such as in-river versus transport mortality that can be measured directly."

In addition to the ISAB's comments and flaws of the upstream/downstream approach that have been identified previously (Zabel and Williams 2000; Williams et al. 2005), we provide two additional comments:

- Weak scientific methodology. The standard scientific method operates by stating a null and alternative hypotheses and considering *all* available information in an effort to reject the null hypothesis. Science does not work by laying out a hypothesis, then saying it is correct unless positive proof exists to show that it is wrong. Yet, this is what has occurred here.

- Ignores data from other systems. Data on natural sockeye salmon populations in Bristol Bay have shown similar trends in overall productivity as have the upstream/downstream comparisons used by CSS. Overall productivity of the Bristol Bay populations increased and decreased over a period of decades, concomitant with major changes in ocean conditions. However, some of these eight closely related populations demonstrated strikingly divergent temporal patterns (Hilborn et al. 2003; Peterman et al. 2003). Yet, the analyses comparing Snake River and John Day River Chinook salmon populations assume that changes in temporal patterns do not exist. The Bristol Bay data suggest a lack of foundation for this assumption.

9. No clear direction exists to argue for continuing the large releases of hatchery fish for the purposes of 'comparative' survival. This is based on: 1) It does not appear that hatchery Chinook salmon provide any useful information related to wild Chinook salmon other than when SARs for hatchery Chinook salmon go way up or way down, proportionately, so do SARs for wild Chinook salmon. This could be determined from a much smaller number of PIT-tagged fish or from adult returns by comparing the clipped to unclipped population. 2) The CSS results indicate that on an annual basis, transportation would benefit hatchery Chinook salmon but not wild Chinook salmon. Since the distribution of hatchery Chinook salmon past lower Granite Dam is much more compressed than that of wild Chinook salmon, it is not clear that even analyses on a temporal basis with hatchery Chinook salmon would provide information on how best to operate the system for wild Chinook salmon. 3) Hatchery Chinook salmon have a wide range in return rates. McCall fish do particularly well, and have a different distribution than Dworshak fish. Which hatchery fish then represent wild fish?

cc: John Stein
John Ferguson
John G. Williams
Bruce Suzumoto

References:

Berggren, T., and coauthors. 2005. Comparative Survival Study (CSS) of Hatchery PIT-tagged Spring/Summer Chinook and PIT-tagged Summer Steelhead; Mark/Recapture Activities and Bootstrap Analysis. Fish Passage Center and Comparative Survival Study Oversight Committee Annual Report to BPA, 2005 Annual Report, Project No. 199602000, 128 electronic pages, (BPA Report DOE/BP-00025634-1) (Available at: http://pisces.bpa.gov/release/documents/documentviewer.aspx?doc=00025634-1)

Copeland, T., J. Johnson, and G. Bunn. 2007. Idaho natural production monitoring and evaluation. Report of Research by Idaho Department of Fish and Game to Bonneville Power Administration, 2005-2006 Annual Report, Project No. 199107300 (BPA Report DOE/BP-00023363-1), 53 electronic p. (Available online at: http://pisces.bpa.gov/release/documents/documentviewer.aspx?doc=00023363-1).

Hilborn, R., T. P. Quinn, D. E. Schindler, and D. E. Rogers. 2003. Biocomplexity and fisheries sustainability. Proceedings of the National Academy of Sciences of the United States of America 100(11):6564-6568.

ISAB (Independent Scientific Advisory Board). 2006. ISAB review of the 2005 Comparative Survival Studies' Annual Report and applicability of Comparative Survival Studies' analysis results. Prepared for Northwest Power and Conservation Council, Columbia River Basin Tribes, and the National Marine Fisheries Service, 2006-3. 21 p. (Available at: http://www.nwcouncil.org/library/isab/isab2006-3.htm).

ISAB (Independent Scientific Advisory Board). 2007. ISAB Latent Mortality Report. Prepared for Northwest Power and Conservation Council, Columbia River Basin Tribes, and the National Marine Fisheries Service, 2007-1. 28 p. (Available at: http://www.nwcouncil.org/library/isab/isab2007-1.htm).

Peterman, R. M., B. J. Pyper, and B. W. MacGregor. 2003. Use of the Kalman filter to reconstruct historical trends in productivity of Bristol Bay sockeye salmon (Oncorhynchus nerka). Canadian Journal of Fisheries and Aquatic Sciences 60:809-824.

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(6):1196-1207.

Williams, J. G., and coauthors. 2005. Effects of the Federal Columbia River Power System on salmon populations. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-63, 150 p. (Available online at http://www.nwfsc.noaa.gov/publications).

Zabel, R. W., and J. G. Williams. 2000. Comments on "Contrasting patterns of productivity and survival rates for stream-type Chinook salmon (Oncorhynchus tshawytscha) populations of the Snake and Columbia rivers" by Schaller et al. (1999). Canadian Journal of Fisheries and Aquatic Sciences 57(8):1739-1741.

FISH PASSAGE CENTER

1827 NE 44th Ave., Suite 240, Portland, OR 97213

Phone: (503) 230-4099 Fax: (503) 230-7559

http://www.fpc.org

e-mail us at fpcstaff@fpc.org

August 31, 2007

Marvin Shutters
Derek Fryer
US Army Corps of Engineers
Walla Walla District
201 N. Third Avenue
Walla Walla, WA 98362-1876

Dear Mr. Shutters and Mr. Fryer:

Thank you for your review of the Draft, Ten-year Retrospective Summary Report. The following response was developed by the Comparative Survival Oversight Committee, (Committee) comprised of, the Columbia River Inter-tribal Fish Commission, the Washington Department of Fish and Wildlife, the Oregon Department of Fish and Wildlife, the Idaho Department of Fish and Game, and the US Fish and Wildlife Service. As you are aware the Comparative Survival Study is a joint project of the agencies and tribes. The study design, the implementation of the study and the analysis are carried out collaboratively among the sponsoring fish and wildlife management agencies. The Committee has developed the following response to your general comments, which are followed by the response to each specific comment.

General Comments

The majority of your comments were presented in a narrative discussion fashion offering broad general ideas and broad alternative philosophies. Recommendations were made regarding how the region should address management issues. Although we found the discussion interesting, the topics you discuss are better addressed in other regional forums. We found it difficult to relate the discussion to the specific aspects of the Draft CSS Ten-year Retrospective Report. We carefully considered the discussion where it was specific to the Ten-year Draft Report and have attached (attachment 1) our specific responses to each individual point in your comments. In response to some of the general discussion, we emphasize that the CSS study uses regionally accepted methodology and analysis, supported in a large body of scientific literature. In addition the CSS study is reviewed annually and the Oversight Committee addresses the regional comments received.

The CSS Oversight Committee is grateful for the significant investment by the COE in the review and preparation of comments on the draft report. The report has been improved as a result of addressing and incorporating comments. We look forward to future positive collaboration with the COE on future CSS monitoring and evaluation.

Sincerely

Michele
Project Leader, Comparative Survival Study

Attachment 1

<u>Reviewer Comment</u> *re: Page 3 line 9: Because the CSS SAR results fail to meet the NPCC SAR objectives, it appears that collecting and transporting juvenile spring/summer Chinook and steelhead at Snake River Dams did not compensate for the effects of the FCRPS on the survival of these fish while migrating through the hydrosystem.*

This statement contains flawed logic. That SAR are lower than objectives provides no evidence that FCRPS related mortality is the reason. Observed SAR are the expression of the total mortality experienced of the sample population. There are many sources of mortality in addition to effects of the FCRPS. A few examples of non-FCRPS related sources of mortality could include: predation, harvest, infection by pathogens, suboptimal environmental conditions, congenital abnormalities, etc. Most of these can occur prior to entry into the hydrosystem, or in the estuary and ocean. Further, it is likely that some of the observed mortality in the hydrosystem is compensatory not additive.

<u>Response:</u> We don't assert that the hydrosystem is the only factor influencing Snake River SARs. However, the other factors that COE cites would also be expected to affect downriver Chinook and steelhead stocks. It turns out that overall SARs from wild Snake River spring/summer Chinook and steelhead clearly have been less than NPCC objectives (minimum 2%, average 4%) across a wide range of ocean and migration conditions, whereas wild stream-type Chinook from downriver populations passing fewer dams have fared much better (see Figure 5.15). Post-Bonneville differential mortality between transported and in-river migrants is differential delayed mortality because it takes place after fish have transited the FCRPS. Moreover, our conclusion that transportation did not fully compensate for FCRPS effects is completely consistent with the NWFSC "Effects memo" (Williams et al. 2005) conclusion (p. xvi) that "transportation is not a panacea for negative effects of dams on fish stocks." We are not sure what the COE statement that some mortality is compensatory rather than additive refers to.

An alternative index for describing and making inferences about the total, overall effect of the hydrosystem on smolt to adult survival is hydrosystem survival. This metric does not refer to absolute values of SAR, yet encapsulates in one quantity everything about the effects of the hydrosystem on smolt survival. It requires estimates of *D*, latent mortality of untransported smolts, and proportion of the migration which is transported. We included this in the review draft of the 10 year report, but have dropped it in the revised draft, due to the need for quantities estimated outside of the CSS. Without either metric (hydrosystem survival or absolute SARs), we would have no way of making inferences about the overall efficacy of transportation-based hydrosystem strategies.

<u>Reviewer Comments</u>: *on 'Estimation of D" page 51 lines 8-11; and all relevant text on pages 58-60; all tables in appendix D-21-D28.*

<u>Response</u>: The statement that "The T/I ratio thus gives us a valid (less biased) comparison of in-river to transportation outmigration life-histories" is not generally true. *TIR* alone is sufficient for comparing *some* management actions to each other. *TIRs* do reflect the overall benefit of

transportation, compared to in-river migration, under the current operation and configuration of the hydrosystem. We estimate and report *TIR*s. However, the overall value of transportation in avoiding jeopardy and promoting recovery depends on hydrosystem survival, which is sensitive to the amount of delayed mortality of both transported and untransported fish. *D* is a frequently used metric that reflects any latent mortality specific to transported fish. See, e.g., Kareiva et al. (2000); Peters and Marmorek (2001); Wilson 2003, Zabel et al. (*in press*).

The claim of bias in *D* due to poor fish condition is a non sequitur. Any culling of weak in-river fish is properly reflected in *D*. High survival in barges due to shielding from mortality that results in later mortality is a consequence of barging, and is properly reflected in *D*. *D* measures the relative survival of transported fish, post-Bonneville, to the survival of untransported fish, post-Bonneville. The reasons for this differential mortality is irrelevant in its estimation. Reasons for *D* being less than one can be postulated; some causes may be addressable but others, such as the shielding of weaker fish from mortality they would otherwise experience leading to those fish dying at a higher rate once they are exposed to estuarine or ocean challenges, are an unavoidable feature of transportation.

Reviewer Comment: Equation 3.9, pg 58, line30 10yr CSS Report Draft

Response: We certainly agree that *TIR* can be > 1 even if *D* is < 1. Nowhere do we claim otherwise. *D* measures something different than *TIR*; we don't make that claim that D < 1 indicates transportation doesn't provide a survival benefit relative to in-river migration in the hydrosystem as currently configured. That's one reason we have *TIR* = 1 lines in Chapter 4 figures, but don't put a *D* = 1 line on the figures showing *D* distributions in Chapter 4. It is unfortunate if this is misunderstood, but we have not promoted this misunderstanding. However, comparison of the observed *D* to 1 is informative about the existence and level of delayed mortality due to transportation, which is useful in modeling and to answer certain questions about the impacts of the hydrosystem. *D* does not "ignore" passage-related mortality; in fact an estimate of such mortality is explicitly required to estimate *D*. In prospective passage modeling, constant (or stationary) *TIR* leads to inflated predictions of transport SAR and hence *D*, if increased in-river survival is modeled, because *D* is directly proportional to *TIR* and an increase in in-river SAR requires a corresponding increase in transport SAR for *TIR* to be constant. Explicitly modeling *D* rather than *TIR* avoids the problem of spuriously increasing post-Bonneville survival of transported fish due to increased in-river survival of untransported fish.

TIRs directly reflect any passage mortality to due poor fish condition (or anything else), and these estimates are presented the report as prominently as *D* estimates. In estimating *D*, we do not need to take into account how many transported fish may be doomed to die after release from barges because of poor condition; we need only a reasonable estimate of the mortality before they are released. See previous comment about culling of in-river fish being properly reflected in *D*. Variation in *D* between years and over the migration season can be and has been addressed, in the CSS report and elsewhere. Any "complications in interpretation" due to variation in *D* are would of course apply to *TIR*, which varies inter-annually and over a migration season as well.

We agree that *TIR* is more direct than "*D*" estimation. But to the extent that casual estimates of in-barge mortality estimated by transportation operators (0.02) is correct, "*D*" provides a second way of evaluating the efficacy of transportation. We need to be mindful, however, that both provide relative estimates of transport effectiveness that may be "moot" (Mundy et al 1994) if absolute survival is insufficient for survival and recovery.

Reviewer Comment: Page 102 line 14: If in-river survivals are similar for C_1 and C_0 groups, as generally assumed, the differential SAR is evidence of delayed mortality for bypassed fish (see Budy et al. 2002).

Response: The detection probability model selection exercise in the 2006 CSS Annual Report (Chapter 9) looked at wild Chinook tagged and released above LGR. The finding was that survival-detection probability model selection provided no clear indication of a biologically meaningful relationship between individual size and detection probability at LGR (or any downstream site). In all cases, size differences between detected and undetected fish, where statistically significant, were less than or equal to 2 mm.

Comments on the CSS Ten-year Retrospective Analysis Report

From: Marvin Shutters and Derek Fryer, Walla Walla District, COE
Date: Submitted to the FPC via e-mail on 27 July 2007.

Page 3 line 9:
Because the CSS SAR results fail to meet the NPCC SAR objectives, it appears that collecting and transporting juvenile spring/summer Chinook and steelhead at Snake River Dams did not compensate for the effects of the FCRPS on the survival of these fish while migrating through the hydrosystem.

This statement contains flawed logic. That SAR are lower than objectives provides no evidence that FCRPS related mortality is the reason. Observed SAR are the expression of the total mortality experienced of the sample population. There are many sources of mortality in addition to effects of the FCRPS. A few examples of non-FCRPS related sources of mortality could include: predation, harvest, infection by pathogens, suboptimal environmental conditions, congenital abnormalities, etc. Most of these can occur prior to entry into the hydrosystem, or in the estuary and ocean. Further, it is likely that some of the observed mortality in the hydrosystem is compensatory not additive.

Reviewer Comments on 'Estimation of D"
page 51 lines 8-11; and all relevant text on pages 58-60; all tables in appendix D-21-D28.

CSS estimates of "D" have assumed a transport-to-release below Bonneville Dam survival rate of 98%. In light of new research data indicating that a high proportion of fish transported are in poor health prior to being collected (Loge et. al 2007), previous estimates of "D" may not reflect the true benefit of transportation. A proportion of the transported fish likely die below Bonneville for reasons unrelated to barging, and yet these mortalities are reflected in the transportation SAR used in the calculation of D. Conversely, the same fish of poor health that remain in-river do not get included into in-river "D" estimates as they likely die prior to passing Bonneville Dam which is the starting point to estimate the in-river SAR used in the calculation of "D". Transport to In-River SAR (TIR) ratios do reflect the true benefit of barging as this comparison includes the poor-health juvenile fish in both the transport and in-river SAR estimates. The T/I ratio thus gives us a valid (less biased) comparison of in-river to transportation outmigration life-histories.

Equation 3.9, pg 58, line30 10yr CSS Report Draft

$$D_1 = (SART_0 * S_r)/(SARC_0/S_t)$$

If we rearrange this equation mathematically, we get $SART_0/ SARC_0 * S_r/ S_t$; which is essentially the TIR equation (pg 58, line 4) multiplied by S_r/ S_t. The CSS assumes a 0.98

S_t which is very close to 100% survival from loading in a barge/truck to release below BON. So the terms in this equation that are heavily influencing the resulting D are the $SART_0$/ $SARC_0$ ratio, and S_r. The CSS Report does mention the importance of the S_r estimate the D estimate (page 58 lines 37-37). If, for a moment, we assume $SART_0$/ $SARC_0$ is very close to 1, then D = S_r if S_t is close to 1 (or .98 from your report). It is clear to see the relationship between D and S_r/ S_t, not D and its relationship to the value 1. Therefore, D-values greater than S_r/ S_t indicate a benefit from transportation. If we look at any of the CSS data tables (D-21-D-28, pages D-14 through D-16) this relationship becomes very clear. If D-values are greater than S_r, TIR values are also generally greater than 1, depending on the $SART_0$/ $SARC_0$ ratio. For example, the first row in table D-22 and (1997 out migration year for PIT-tagged Rapid River Hatchery spring Chinook salmon) shows a S_r of 0.33, ad TIR of 1.73 and a D of 0.61. This TIR indicates that transportation resulted in a 73% higher return rate of adults than outmigrants with an in-river life history. Moreover, if we look at the geometric mean values from this same data table (D-22) results show an S_r of 0.52, a TIR of 1.46, and a D of 0.81, there was a 46% higher return rate of transported fish relative to in-river fish over a 7 year time period. All of the data tables (D-21 through D28) follow the same pattern; as D is greater than S_r , TIR is greater than1 indicating a transportation benefit. As the $SART_0$/ $SARC_0$ ratio becomes smaller, the D-value only needs to be larger than the S_r (compared to when the $SART_0$/ $SARC_0$ is close to 1) to indicate a survival benefit from transportation. The S_r , TIR, and D data tables clearly demonstrate that D-values can be significantly less than 1 yet the TIR is over 1 (e.g. transportation benefit) a long as D is greater than the S_r.

The calculation of D was created in the PATH process in order to improve modeling efforts to understand the difference in survival between transported and in-river migrating juvenile salmonids below Bonneville dam. To our knowledge, D was not intended to be a management index of transportation benefit/non benefit. However, this value has been misinterpreted throughout the region, because of the assumption that D-values lower or higher than 1 indicate a transportation non-benefit or benefit, respectively. D-values less than 1 indicate there is a differential survival rate below Bonneville Dam between in-river and transported groups, but D-values less than 1 DO NOT indicate a non-benefit from transportation. If D is viewed in relation to S_r or the in-river hydosystem survival estimate, then the D-value takes on a more relevant meaning: $D < S_r$ = no transport benefit, D greater than S_r = transport benefit.

The TIR ratio is a more unbiased metric for evaluating the benefits of transportation because it takes into account the mortality of in-river migrants and subsequent survival of barged fish once collected a LGR. The D calculation removes the unhealthy fish from the In-river fish calculations as most, if not all are culled before reaching Bonneville Dam, but does not remove them from the Transport category. We believe there is significant proportion of unhealthy fish that are transported that die shortly after release from the barge/truck (20-50% of transported hatchery spring Chinook exhibit mortality shortly after release, Loge et al. 2007). Had these unhealthy fish been returned to migrate in-river, their fate would be the same (or potentially worse) than had they been barged.

Moreover, using the D-metric ignores the passage related mortality transportation is undertaking to avoid.

To state this another way, if we knew that 20-50% of the barged fish were going to die, (likely from infectious disease present prior to being collected and transported) shortly after release from the barge/truck below BON, wouldn't we change the S_t to more accurately reflect this? Instead of using S_t = 98%, shouldn't we use a 50-80% S_t value if we could accurately estimate it. How would this change all the D-values? More importantly, how would we accurately estimate the percent of unhealthy fish that are barged from year to year? **If you view D relative to the S_r you might get a better estimate of post BON survival of the healthy fish that are barged** and not necessarily need to develop estimates of the proportions of unhealthy fish that may die shortly after release from the barge/truck. D-values are complicated to interpret because the $SART_0$/ $SARC_0$ ratios and annual S_r estimates change over the season and from year to year. Further complicating an interpretation of D is the potential violation of the assumption that S_t is 0.98, a value that is likely much less than 0.98, and a value that is likely to vary from year to year.

In conclusion, interpretation of benefits of transportation should be made using the TIR ratios, which provide a valid metric of in-river and transportation survival benefits. D-values should not be used as an index of transportation benefits as it relates to 1, but as it relates to S_r (or more accurately S_r/ S_t). D is one of the most complicated and controversial subjects within Snake and Columbia River Basin. This topic needs much more discussion in future and current drafts of CSS reports so that results of life cycle PIT tag studies are correctly interpreted.

Page 102 line 14:
If in-river survivals are similar for C1 and C0 groups, as generally assumed, the 15 differential SAR is evidence of delayed mortality for bypassed fish (see Budy et al. 2002).

Another potential explanation for the observed difference in SAR for C1 and C0 is the documented size selectivity of most bypass systems. The C0, or uncollected fish tend to be larger. Larger fish would also, be expected to have higher SAR.

References Cited

Loge, Frank, Joseph Dietrich, Deborah Boylen, Dina Spangenburg, Claudia Bravo, Donald Thompson, Erik Loboschefsky, Mary Arkoosh, and Tracy Collier. Disease Susceptibility of Hatchery-Reared Yearling Snake River Spring/Summer Chinook Salmon with Different Migration Histories in the Columbia River. Draft report Submitted to the Army Corps of Engineers, Walla Walla District. March 2007.

FISH PASSAGE CENTER

1827 NE 44th Ave., Suite 240, Portland, OR 97213
Phone: (503) 230-4099 Fax: (503) 230-7559
http://www.fpc.org
e-mail us at fpcstaff@fpc.org

August 31, 2007

Dr. James Anderson
University of Washington
Professor, School of Aquatic and Fishery Sciences
1325 -4th Ave., Suite 1820
Seattle, WA 98101

Dear Dr. Anderson:

Thank you for your review of the Draft, Ten Year Retrospective Summary Report. The following response was developed by the Comparative Survival Study Oversight Committee, (Committee) comprised of, the Columbia River Inter-tribal Fish Commission, the Washington Department of Fish and Wildlife, the Oregon Department of Fish and Wildlife, the Idaho Department of Fish and Game, and the US Fish and Wildlife Service. As you are aware the Comparative Survival Study is a joint project of the agencies and tribes. The study design, the implementation of the study and the analysis are carried out collaboratively among the sponsoring fish and wildlife management agencies. The Committee has developed the attached response (attachment 1) to your comments.

Sincerely

Michele Dehart
Project leader, Comparative survival Study

Attachment 1

Reviewer Comment: Result using S were not presented.

Response: Results on the model fits (AIC values) using S as dependent variables were presented in Table 2.2 and the variables that were selected were reported on page 23 of the draft report. The revised version contains a table describing the models that were fit with S as the dependent variable, the parameter estimates for the best-fit model, and an expanded comparison of the approach of modeling S versus modeling instantaneous mortality rates for all reaches as species groups evaluated.

Reviewer Comment: Mathematically the analysis based on Z is not valid.

Response: We believe that we are on firm ground mathematically with the use of Z. The mathematics of instantaneous mortality (*Z*) go back to Malthus (1798). The exponential law of mortality, which is based on *Z*, has been called the "first principle" or "first law" of population dynamics (Turchin 2003). The formula we used for estimating *Z* is the maximum-likelihood estimator for *Z* (Seber 1982, p. 216). The exponential law of mortality forms the basis for nearly all fisheries population dynamics models (Quinn and Deriso 1999).

Reviewer Comment: The analysis and conclusions based on Z should be deleted from the report and replaced with the analysis based on S.

Response: We provide a comparison of three approaches for predicting survival rates, including one that uses *S* as the dependent variable. By nearly all performance measures, the approach based on *Z* outperformed the analyses that used *S* as the dependent variable.

Reviewer Comment: The mathematical error in their analysis can be demonstrated as follows. Z contains information on fish travel time fft since it is defined

$$Z = -\frac{\log S}{ftt} \qquad (1)$$

Response: We do not disagree that *Z* reflects changes in FTT (the denominator). However, *Z* also reflects changes in survival (the numerator). We found that most of the variation in *Z* was associated with variation in *S* (49-58%), whereas only a small amount of the variation in *Z* was associated with variation in FTT (2-13%).

Reviewer Comment: *However, fish travel time decreases with increasing Julian day and water travel time. This has been established in earlier studies (Zabel et al. 1997, 1998, in press). The CSS study found a similar result*

$$\log ftt = a_0 - a_1(ju) + a_3(ju)^2 + a_4(wt) - a_5(wt)^2 \quad (2)$$

Response: We find it peculiar that you have chosen to omit the spill variables that we reported from your mischaracterization of our work. Recall, if you will, that spill was found to reduce fish travel time for all species and all reaches analyzed. We do not disagree that Julian day and water transit time also affect fish travel time. However, we clearly demonstrated that the average percent spill was a primary determinant of fish travel time, with higher levels of spill associated with reductions in fish travel time.

Reviewer Comment: *Therefore, Z is a function of ju and wt independent of any effect of these variables on S.*

Response: As noted above, most of the variation in *Z* is associated with variation in survival (49-58%), whereas only a small amount of the variation in *Z* was associated with variation in *FTT* (2-13%). Given these results, and the fact that *Z* is calculated as a function of survival and fish travel time (essentially averaging total mortality over a period of time), it is unclear what your basis is for arguing that *Z* is independent of *S*.

Reviewer Comment: *In fact, Zabel et al. in press analyzed the effects of similar covariates on survival (S) and found temperature was a dominant factor in the upper reach and the only factor in the lower reach. These results stand in variance to the claims in the CSS report (lines 3-9 page 24)*

Response: The quote you refer to has nothing to do with modeling the effects of covariates on *S*, temperature or otherwise. The quote summarizes the instantaneous mortality rates that were observed in the upper and lower reaches and what the relative magnitude of those values mean.

Reviewer Comment: *The claim is not supportable. In the lower reach, mortality is independent of time in reach (Zabel et al in press). Mortality depends on temperature so the results in the CSS study reflect the effect of wt and ju on fish travel time, not on survival.*

Response: Again, the quote you refer to has nothing to do with modeling the effects of covariates on *S*, temperature or otherwise. Rather, it simply summarizes the data. See above response.

Reviewer Comment: *Relating river conditions to Z, and not S, does not reveal the effect of temperature on survival, contrary to the claims in the CSS report. The report states (line 17-19 page 24)*

Response: We did not find that temperature was an important factor for explaining patterns of variation in instantaneous mortality rates, survival rates, or fish travel times. Only in the lower reach for steelhead (the data set with the greatest level of imprecision) was temperature identified as being associated with instantaneous mortality rates. Because temperature did not explain variation in the data in the upper reach where the data were more precise, we suggested that the identification of temperature as a primary determinant of instantaneous mortality rates for steelhead *may* be a spurious correlation. However, if you had continued to read the draft report, you would have read that we offered the alternative explanation that the factors influencing mortality rates in the lower reach (i.e., temperature) may be different than those operating in the upper reach.

Zabel et al. (in press) found temperature was important in the upper reach. Furthermore, the 2001 data reveals a strong temperature effect not a flow effect (Anderson 2003). In 2001, flow increased and decreased over the migration season while survival dropped steadily (Figure 1). However, survival dropped as temperature increased showing (Figure 2). The CSS model is incapable of capturing this pattern.

A visual inspection of the predicted survival rates in Figure 2.9 of the draft report clearly demonstrates that the model developed by the CSS is quite capable of capturing the pattern of survival in 2001, as well as the other years analyzed, for both Chinook and steelhead.

Comments on Chapter 2 of Comparative survival study draft 5/30/2007

James Anderson
Professor, School of aquatic and Fishery Sciences
University of Washington
June 29, 2007

Conclusions and Recommendations

In Chapter 2 of the CSS Ten-year Retrospective Analysis Report the effects of environmental variables on fish passage survival were analyzed using survival (S) and instantaneous mortality (Z). The report draws conclusions based on the analysis using Z. Result using S were not presented. Mathematically the analysis based on Z is not valid. The analysis and conclusions based on Z should be deleted from the report and replaced with the analysis based on S.

The authors model the instantaneous survival (Z) and survival (S) as a function of water travel time (wt), Julian day (ju), temperature (te), turbidity (tu) and spill (sp). However, survival results are only discussed for the analysis with Z.

The equation selected is

$$Z = a + b * wt + c * wt * ju \quad (1)$$

where instantaneous mortality increases with water travel time and Julian day.

From this analysis, the report states that (lines 7-11 page 23)

> "The models for characterizing instantaneous mortality rates provide information on how and why mortality rates may vary (Figure 2.17). For wild Chinook in the LGR-MCN reach, instantaneous mortality rates are estimated to remain low throughout the season when water transit times are short (5-d). As water transit times get longer, instantaneous mortality rates rise rapidly over the season."

This result is problematic and misleading because Z is related to wt and ju whether or not survival is related to these variables. The important issue involves what affects survival not instantaneous mortality which can change by travel time without a change in survival.

The mathematical error in their analysis can be demonstrated as follows. Z contains information on fish travel time *ftt* since it is defined

$$Z = -\frac{\log S}{ftt} \quad (2)$$

However, fish travel time decreases with increasing Julian day and water travel time. This has been established in earlier studies (Zabel et al. 1997, 1998, in press). The CSS study found a similar result

$$\log ftt = a_0 - a_1 (ju) + a_3 (ju)^2 + a_4 (wt) - a_5 (wt)^2 \qquad (3)$$

Therefore, *Z is a function of ju and wt independent of any effect of these variables on S.*

It follows, that effect of wt in equation (1) is strongly condition by its effects on ftt in equation (3). When using Z as the dependent variable it is not possible resolve the effect of wt on survival. In fact, Zabel et al. in press analyzed the effects of similar covariates on survival (S) and found temperature was a dominant factor in the upper reach and the only factor in the lower reach. These results stand in variance to the claims in the CSS report (lines 3-9 page 24)

> "Several patterns have emerged from the examination of instantaneous mortality rates. First, for both species, instantaneous mortality rates in the MCN-BON reach are roughly double those in the LGR-MCN reach (Table 2.3). This means that one additional day spent in the lower reach will result in twice the level of mortality that would occur with an additional day spent in the upper reach."

The claim is not supportable. In the lower reach, mortality is independent of time in reach (Zabel et al in press). Mortality depends on temperature so the results in the CSS study reflect the effect of wt and ju on fish travel time, not on survival.

Relating river conditions to Z, and not S, does not reveal the effect of temperature on survival, contrary to the claims in the CSS report. The report states (line 17-19 page 24)

> "Given that temperature was not identified as a primary factor in the upper reach where the data were more precise, the identification of temperature in the lower reach as a primary determinant of instantaneous mortality rates in steelhead may be a spurious correlation."

Zabel et al. (in press) found temperature was important in the upper reach. Furthermore, the 2001 data reveals a strong temperature effect not a flow effect (Anderson 2003). In 2001, flow increased and decreased over the migration season while survival dropped steadily (Figure 1). However, survival dropped as temperature increased showing (Figure 2). The CSS model is incapable of capturing this pattern.

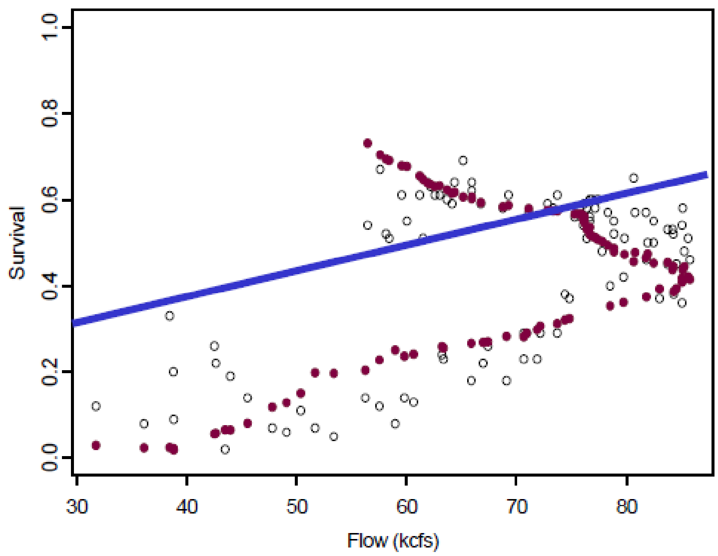


Figure 1. Spring chinook survival vs. flow between Lower Dam and McNary dam for 2001. Survival estimated with designated (○) survival estimated with the CBR model designated Line depicts the low flow segment of NOAA's hockey stick flow/survival relationship (from Anderson 2003a).

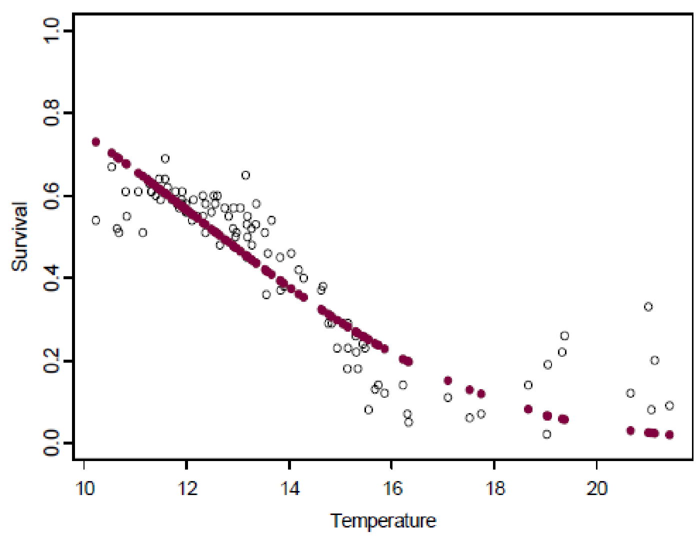


Figure 2. CBR model showing relationship between chinook survival and temperature over the reach LGR and MCN in 2001. Survival estimated with PIT tags designated (○) survival estimated with the CBR model designated (●) (from Anderson 2003a).

References

Anderson JJ. 2003 Toward a Resolution of the Flow/Survival Debate and the Impacts of Flow Augmentation and Water Withdrawal in the Columbia/Snake River System. www.cbr.washington.edu/papers/jim/towards_res.pdf

Anderson J.J. E. Gurarie and R. W. Zabel (2005). Mean free-path length theory of predator-prey interactions: application to juvenile salmon migration. *Ecological Modelling* 186:196-211.

Zabel, R. W. J. Faulkner, S.G. Smith, J. J. Anderson C. Van Holmes, N. Beer, s. Iltis, J. Krinkie, G. Fredicks, B. Bellerud, J. Sweet and A. Giorgi. (2007). Comprehensive Passage (COMPASS) Model: a model of downstream migration and survival of juvenile salmonids through a hydropower system. Hydrobiologia.

Zabel, R.W., J.J. Anderson, and P.A. Shaw. 1998. A multiple reach model describing the migratory behavior of Snake River yearling chinook salmon (Oncorhynchus tshawytscha). *Canadian Journal of Fisheries and Aquatic Sciences*: 55:658-667.

Zabel, R. and J.J. Anderson. 1997. A model of the travel time of migrating juvenile salmon, with an application to Snake River spring chinook salmon. *North American Journal of Fisheries Management*, 17:93-100.