Memorandum (ISAB 2010-5) October 14, 2010

To: ISAB Administrative Oversight Panel
    Bruce Measure, Chair, Northwest Power and Conservation Council
    Paul Lumley, Executive Director, Columbia River Inter-Tribal Fish Commission
    Usha Varanasi, Science Director, NOAA-Fisheries Northwest Fisheries Science Center

From: Rich Alldredge, ISAB Vice Chair

Subject: Review of the Comparative Survival Study 2010 Draft Annual Report to suggest topics for further review and to provide comments to improve the Annual Report

Background

The Northwest Power and Conservation Council’s 2009 amendments to the Columbia River Basin Fish and Wildlife Program call for the continuation of the fish passage related functions currently conducted by the Fish Passage Center (FPC). The primary functions are to provide technical assistance and information to fish and wildlife agencies in particular, and to the public in general, on matters related to water management, spill, and other passage measures. The Program also calls for the Fish Passage Center’s Oversight Board to ensure that the functions are implemented consistent with the Program. To do this, the Program specifies that the Oversight Board will work with the Center and the Independent Scientific Advisory Board (ISAB) to organize a regular system of independent and timely science reviews of the Center’s analytical products.

The Oversight Board, ISAB, and FPC director are currently establishing guidelines for this regular review. Although the guidelines are not finalized, the Oversight Board agreed that the ISAB should implement the first review assignment item included in the draft guidelines. Item 1 specifies that a subgroup of the ISAB will “initiate an examination of the FPC and CSS draft annual reports when these reports are released for public comment. As part of the examination, the subgroup will look at the annual reports to ensure that work products, methodologies, and analyses appropriate for potential science review have been considered.”

The draft guidelines also include criteria for identifying FPC analyses/products for ISAB review. These include the introduction of new or novel analyses; new conditions or data bring old
analyses into question; and/or consensus cannot be reached in the region on the science involved in the product.

In July 2010, the ISAB conducted its first review under the draft guidelines and reviewed the FPC 2009 Draft Annual Report. That review was completed within the FPC’s public comment period for that report. The ISAB provided detailed comments on the draft report and identified three FPC analyses that met the criteria for further review (ISAB 2010-4). Because the guidelines for ISAB review of FPC products are not finalized, the ISAB is waiting to begin a review of those three analyses until the FPC Oversight Board completes and communicates its consideration of the suggested reviews. In the interim, the Comparative Survival Study (CSS) 2010 Draft Annual Report was released for public comment until October 15, 2010. The ISAB’s comments on the draft report follow below.

Summary

The ISAB once again acknowledges the continuing improvement in the organization, clarity, and writing quality of recent annual reports, as exemplified in the 2010 CSS Annual Report. This ISAB review begins by providing general comments on the content of the 2010 CSS Annual Report and then follows with specific editorial suggestions. In addition to providing these comments intended to improve the final annual report, the ISAB recommends further independent peer review of the methods proposed in Chapter 7 for modeling bypass systems. This modeling effort introduces new or novel analytical approaches and thus meets the draft guidelines criteria for more in-depth review.

1. General Review Comments

Overall, the presentation is well organized and well refined. An overarching comment is that connections with larger ecological concerns are not apparent. That is, there appear to be opportunities to involve researchers working on studies of other species, food webs, physiology, contaminants, and disease. Such combined studies might give added insights into mechanisms causing the observed temporal patterns in migration and survival.

Inclusion of a list of acronyms would help many readers.

Chapter 1

The description of how the differently marked cohorts are used to translate into SARs and TIRs is valuable and much appreciated. Including a graphic/schematic would make it easier for the reader to understand this complex material. The use of differently marked cohorts reappears in Chapter 4 and the need for visuals is even greater there. This is an important procedural change, and clarity could be improved by including a general visual here and perhaps two to three more visuals in Chapter 4, where the detail is mathematically dense.
Figures 1.2 and 1.3 could be made clearer by highlighting Release and Detection sites. The figures contain much detail so the main points of the figures may not be obvious. Table 1.1 is very instructive. It would also be useful if the numbers of fish used in the CSS were noted as an approximate fraction of releases from the basin and from above Lower Granite Dam (LGR). That is, what percentage of all PIT tag releases above LGR is due to CSS?

If pre-release PIT tag loss is monitored at hatcheries (page 15) how is PIT tag loss assessed post-release? It is not clear how much design work goes into pre-assigning stocks to be PIT tagged. CSS appears to play a major role, but it is not clear if this is a formal or ad hoc role.

Chapter 2

The figures and tables in Chapter 2 are clear and informative.

It may be beyond the scope of this chapter to include statistical analyses, but many of the stated results could be subjected to statistical tests. For example, more specific conclusions could be made as to the actual trends observed such as in 2008 and 2009 compared with the long-term average, between hatchery and wild, and among release groups, rather than a qualitative assessment of the trends. The discussion in chapter 2 is appropriately brief if no statistical tests are to be developed for this chapter. It is not clear whether inclusion of statistical testing has been considered and rejected in this chapter.

Chapter 3

This chapter clearly explains the methods used to estimate juvenile travel times, survival rates, and instantaneous mortality rates and relates these metrics (less clearly) to inter-annual changes in river conditions. Excellent graphics summarize changes in these indices across the 12-year period from 1998 to 2009. The value of this long-term study for addressing complex and ever-evolving questions regarding the effects of the hydropower system on juvenile and adult salmonids is evident in this chapter.

The results of the modeling efforts described in this chapter appear to be successful. This chapter reports on modeling the effects of five variables – average percentage spill, water temperature, water transit time, Julian date, and number of surface passage structures installed – on fish travel times (FTT) and instantaneous mortality rates (Z). Modeled values for FTT and Z in different years and two reaches (LGR to McNary Dam [MCN] and MCN to Bonneville Dam [BON]) were then used to predict survival rates, and predicted survival rates were compared with observed (estimated) survival rates.

Akaike’s Information Criterion (AIC) was used to evaluate model fit and select best-fit models. The report states (p. 57) that “the best fitting models (based on AICc) for mean FTT, [and also
for mean Z, reported in the following paragraph] consistently had model forms with Julian day, water transit time and spill.” The number of surface passage structures was also included in some models. These findings are incompletely described, and thus the basis for selecting best-fit models is unclear. AICc, delta AICc, R², and P-values for alternative models are not reported, nor are the results of model averaging. The relative contributions of the different variables to model fit are not shown. This information may have been omitted for the sake of brevity, but should be reported in future years as the database continues to grow.

Predicted values for Z and FTT in the LGR-BON reach deviated considerably from observed values in some years (Fig 3.1, wild and hatchery Chinook salmon 2008-2009; Fig. 3.3, steelhead, aggregate hatchery and wild, 2004-2009). Discussion of hypotheses that might help explain these instances would be useful to provide context for the results.

Chapter 4

This chapter focuses on the primary objective of the CSS to estimate and compare smolt-to-adult return rates of juveniles passing LGR and returning to LGR as adults and post-hydrosystem survival rates of juveniles passing BON and returning to LGR for both in-river migrating and transported juvenile salmonids. The narrative is extremely careful and explicit with much attention paid to ensuring fair comparisons and to describing the difficulty of what was done to ensure fair comparisons. It would be useful to include flow charts to assist the reader through the adjustments and record keeping described in the narrative.

Standard tagging and detection practice has changed over the years, which has both allowed and necessitated changes in the way various constructs are calculated. Statistical methods for the CSS study have been modified and refined over time and are adequately described in this chapter on pages 67 through 75. SARs, TIRs, and D-values are reported (pages 76-92) in detailed tables and graphs for all study years, as in past annual reports. It would be useful to summarize how to compare these redefined constructs/estimates over time and changing procedures with earlier estimates. Were the earlier estimates approximately correct or systematically off? The procedures, and the attendant estimation methods, will continue to improve, but we will always need to be able to compare over time and across methods. An evaluation of the current state of progress is needed.

Much of the detailed data previously included in the appendices of the CSS reports is now available only on the FPC website. This seems to be a good idea, but the data accessible on the FPC site has not been updated to include the two most recent years of completed adult returns.

The very useful summary at the end of the chapter conveys the larger perspective quite well.
Chapter 5

The introduction effectively describes the nature of the questions to be answered and why they were important. This introduction could serve as a model of an excellent, lucid overview of chapter contents.

Overall the chapter is well written and well thought out. The figures, for example Figure 5.1, were very useful.

It is agreed that, as stated in the methods, “the use of fish detected upstream that were not detected at BON to estimate BON efficiency was the best available measurement of this parameter.” As the authors noted, “this nominal estimator of efficiency could have been inaccurate if fish passed BON undetected and through straying/harvest/mortality were never again detected. This problem was alleviated by comparing these two rates in a fraction (e.g., Success(TX or TO)/Success(Cx). The assumption here was that the rate for passing BON undetected and never being detected again was the same for the transported and in-river fish. Since the fish were from the same species/hatchery, the assumption seemed reasonable.” However, although the approach used may be the best available, it may not be a very good assumption. Transported fish may be more likely to stray into intervening tributaries and not be detected again.

It is also not clear in this section how fall-backs are handled (i.e., those adults that pass upriver and subsequently fall back downriver). They may have been dealt with adequately, but some clarification of how those fish were handled would be useful in this section. Is there evidence that transported fish fall-back more or less frequently than in-river fish?

In the introduction to this chapter, the point was made that best practice is to tag the smolt at the source, rather than at LGR, because of losses of LGR-tagged fish. That also means we should be able to determine the survival rate from the source (with some allowance for handling problems and initial mortality) to LGR. That is not reported here. It also means that we have the potential (or at least the need) to examine return success from LGR, back to source, though detection capability at (or closer to) the source would be needed. It would be useful in future efforts to obtain good estimates of the source to LGR and LGR to the source components.

Chapter 6

This chapter is generally well written. The results provide an informative summary of the long term SARs. It also provides a useful comparison and discussion of differences in SARs calculated through run reconstruction and the CSS PIT tag project, as suggested by ISAB/ISRP review. This discussion is important because of the potential bias in both methods. Further exploration and resolution will be necessary before the full utility of the information is clear; some clarification and care in this section could be a useful step in that direction.
On page 130 L. 2-4 the report states that “it was unclear whether a bias existed in either the run reconstruction or PIT tag SARs due, in part, to uncertainties and assumptions in the run reconstruction methods.” Clarifying that both methods, not just RR methods, are subject to uncertainties and assumptions, would help. In L. 15-17 the report states “…SARs did not appear to be predominantly due to differences in juvenile abundance estimation methods” and cites Tuomikoski et al., 2009 on “similarities” in abundance estimates used in both cases. It would be useful to at least briefly consider the nature of the “similarities” and the uncertainty in both estimates to demonstrate the magnitude of the potential bias from this source.

On Page 131 L.5-23 and page 131 L. 24-46 and page 132 L. 1-38 the report outlines and explores the possible sources of bias in both methods. Although these are very important observations, the discussion of potential bias is much more fully developed in the run reconstruction section than the PIT tag section. Comparable development that sheds some light on the potential magnitude of different sources, or possible methods for exploring, the PIT tag bias would be very useful and would help balance the perspectives. By example in the discussion the potential bias due to non-representative marking of steelhead is invoked as a possible explanation for a lack of correlation between wild and hatchery stocks. Those observations could be used to consider the nature of that source of bias in this section.

Chapter 7

The overall objective for this chapter is to improve understanding of the effects of bypass systems on Chinook salmon and steelhead. Three sets of analyses were conducted:

1) Evaluation of the effects of bypass systems on fish travel time from Lower Granite Dam to Bonneville Dam for fish that are detected in the bypass systems at Little Goose, Lower Monumental, McNary or John Day dams relative to fish that are not detected at those dams.

2) Evaluation of the effects of bypass history on SARs from Bonneville Dam outmigration to Bonneville Dam. Within this analysis two questions were explored. The first was whether multiple bypass experiences negatively affected post-Bonneville Dam SARs. The second question was whether there was evidence that bypass at particular dams was more harmful than others.

3) Evaluation of whether the cumulative effects of bypassing smolts at dams results in increased mortality expressed in SARs by using a random effects meta-analysis.

The chapter contains much detailed information about methods for modeling bypass systems. That said, there is much supporting and background material that has not been included. A more complete review of this chapter is warranted, perhaps in combination with an overall examination of the regional bypass evaluation efforts. The caveats stated in this chapter that caution should be used when interpreting route-specific survival estimates and SARs conditioned on bypass detection provide support for a careful, independent review of the complexity of the data, modeling, and interpretation of results.
In Table 7.2 it may be useful to show non-significant estimates as well to provide an overview of general tendencies.

The use of Box-Cox to select a transformation to reduce the influence of right skewed distributions should be reconsidered. The Box-Cox method is primarily designed to identify transformations that can address heterogeneity of variances, not problems with normality. Using the reciprocal transformation of FTT in Chapter 7 while using the more conventional log transformation of FTT in Chapter 3 deserves justification.

**Miscellaneous editorial comments**

p. 6, line 7. “…long term dataset of annual estimates of the survival rate of annual generations of salmon…”

p. 6, line 12. presumably not just Chinook salmon?

p. 6, line 27. It's should be It's.

p. 6, line 34. “this the 2010 annual report…”

p. 6, line 38. “This current The 2010 annual report…”

p. 6, line 42. “This report completes the includes complete return data (i.e.,3-salt returns) from smolt migration year 2007…”

p. 6, line 42. (ISAB 2010) should be ISAB (2010)

p. 7, line 2. Suggest omitting "...are few, but..."

p. 7, lines 7-10. awkward sentence “Reductions in the number…”

p. 7, line 15. The caption for Figure 1.1 should define the acronyms used in the figure: SAR, TIR, S/S, AND R/S. (The acronyms S/S AND R/S appear unnecessary for this figure).

p. 9, Figure 1.3: Change caption to explain how this figure is different from the previous figure (appears to be a detailed version of just a subset of the Columbia River basin), and perhaps, why it has been included.

p. 10, line 1- DPS/ESU —what is DPS
p. 10, line 4. “…the caveat that we have presented our presentations of Snake River stocks only above do not include stocks below Lower Granite Dam. Also,…as a Mid Columbia group, this was for simplification as this was partly for simplicity, as it is upstream of Bonneville Dam.”

p. 10, line 22. “For hatchery fish, low abundance was a concern as …”. Meaning is unclear.

p. 10, line 33. “…, hatchery-reared fish are tagged…”

p. 10, line 37. “Recapture information can be collected without sacrificing each fish, and lower impacts due to trapping and handling occur where automated detection stations exist reduce impacts from trapping and handling.”

p. 10, line 39. “The Columbia and Snake River Mainstem PIT-tag detectors at the mainstem dams in the Columbia and Snake rivers now allow…”

p. 10, lines 42-46. Awkward sentence should be rewritten.

p. 12, line 2. “…information. This method is used to estimate survival of the total number of fish estimated to approaching…”

p. 12, lines 4-5. Meaning is unclear

p. 12, line 19. (change) CSS is also examines (to) CSS also examines

p. 13, line 14. “Estimates of D isolates excludes mortality occurring during juvenile salmon passage …..dams from and captures any differences in mortality between transported smolts and in-river migrants that occurring afterwards (during time downstream of Bonneville Dam, in the ocean and upon returning upriver as adults to Lower Granite Dam for transported smolts).”

p. 13, line 17. “When D=1 is equal to one it indicates that there is no difference…” and similarly in subsequent sentences.

p. 13, line 28. “1. tagged fish that are detected and collected at Snake...LMN), and transported downstream of BON.”

p. 13, line 39. affects affected

p. 14, line 37. Should explain (at least conceptually) how the number of C0 smolts will be estimated

p. 15, line 19. (change) data is (to) data are

p. 15, lines 40-42. Not a complete sentence
p. 16, line 6. “...operation to rearing only stocks endemic to the Grande Ronde River basin endemic stocks.”

p. 16, line 17. “In coming years With the greater coverage...separation of metrics...by basins should be possible in coming years.

p. 16, line 28. “Pre-release Tag loss and mortality of PIT-tagged fish were monitored before release, and tagging...”

p. 16, lines 30-34. Does not make grammatical sense, as stated.

p. 17, line 1: “…to affect cost savings and avoid redundancy and save costs as ...”

p. 17, line 15: “…Fish and Wildlife; ...”

Tables 1.2 and 1.3: Define “Ch.” and “St.” in caption. Also, why are numbers aggregated for Ch./St.? Would it not be better to present the numbers separately for each species?

p. 20, line 7 “…had lower spill.” Lower than what?

p. 20, line 8. “particularly unique”. Unique is not a word that can be qualified. “Highly unusual” is more appropriate and more scientifically accurate.

p. 20, line 16. delete second “also”

p. 20, line 18. The sentence “The percentage of smolts transported in 2001, 2004, and 2005 were three years with the highest transportation percentages of CSS PIT-tagged wild fish.” is grammatically incorrect and should be rewritten.

Figure 1.4. shows more than just the “average daily values” (title of chart). Presumably the box plots summarize the distribution of daily values, and a brief explanation in the caption seems warranted. Otherwise this is a very effective figure.

p. 47. The statement that survival estimates are not stable should be clarified, perhaps by noting that estimated survival rates exceed 1.

p. 52, lines 12-14. state “Due to sufficient numbers of PIT-tagged hatchery and wild yearling Chinook available, analyses in the LGR-MCN reach were conducted separately for hatchery and wild yearling Chinook....Analyses on the MCN-BON reach included hatchery and wild yearling Chinook and steelhead from the Snake River....”

Comments: (1) This reads as if separate analyses were done for hatchery and wild steelhead (as well as for hatchery and wild Chinook) in both the MCN-BON reach, but this is probably not the intended meaning. (2) In the following paragraphs on p. 52 and p. 53 it is made clear that
hatchery and wild Chinook were in fact combined for analyses of fish travel time, survival, and instantaneous mortality in the MCN-BON reach, apparently contradicting the statement above (in lines 12-14). Minor rewording is needed to correct these problems.

p. 70, lines 33-35: An assessment of how divergent \( C_0 \) and \( C_1 \) are under well estimated situations is needed along with a sense of the reliability of this and other combined estimates.

p. 71, line 2. Insert word “be”: “…transport SARs have tended to [be] low (NOAA 2008).”

p. 73, lines 5-7: The reader needs a rough quantitative sense of how limited the contribution of jacks and mini-jacks is to spawning. The issue may become more relevant if the “jacking rate” increases, due to adaptive pressures in favor of early return.

p. 74, lines 2-6: Confidence bands on any ratio are inevitably asymmetric. The material presented here is not clear in describing precisely how confidence bands were constructed. There are several possible alternatives. More explanation should be provided for clarity.

p. 76. Captions for Figure 4.1, Table 4.1 (and other figures and tables in this chapter) include the sentence “The transport SAR (\( T_x \)) after 2005 is a partial year metric because of the delay in transportation start date.”

Comment: The delayed start to the transportation season after 2005 is adequately explained elsewhere in the text (in several places) and doesn’t need to be repeated in figure and table captions. In addition, the phrase “partial year metric” is not an aid to understanding.

p. 79. The caption for Table 4.2 should be bolded to match the captions for other tables.

p. 88, line 4. Here (and elsewhere?) it would be best to describe the “break even” TIR to one decimal point rather than as an integer: i.e., as 1.0 rather than 1.

p. 96, line 31. A verbal sense of the Chinook jack return (and its likely impact) would be good here, just for perspective.

p. 97, lines 9-10. A verbal sense of the ‘near’ equivalence of \( C_0 \) and \( C_1 \) would be good here, given that they will be averaged. It would also be good to provide some numeric sense of the range of \( BON_{\text{efficiency}} \) here.

p. 97, lines 19-21. Given accumulating information to the effect that transported fish are more prone to stray than are in-river migrants, is this a reasonable assumption? Straying and non-detection are not the same thing, but are straying fish as likely to remain undetected as those that do not stray?
p. 97, lines 30-32. The material presented here is not clear in describing precisely how confidence bands were constructed and there are several possible alternatives. More explanation should be provided for clarity.

p. 99, lines 1-2. Why a t-test, particularly, given the logit modeling?


p. 104, lines 12, 22, 29. These sentences are difficult to interpret. Perhaps a revision such as: “. . . models in the set for wild Chinook was 100% (transport), 100% (spill), 47% (LGR. marking), 25% (temperature), respectively.”

There appears to be one more number than categories.

p. 104, lines 42-46. The text is grammatically awkward and would benefit from revision.

p. 105, line 38. “drop-out rate”

p. 117, line 12. “achieving overall SARs (including jacks) in the…”

p. 117, lines 25-26: “...take actions to improve the likelihood that Columbia River salmon can survive varying ocean conditions.” Is the intention to take actions to improve ocean survival based on predictions of ocean conditions that will be experienced or to compensate for poor survival after the fact by reducing mortality during the return migration (e.g., from fishing, passage)?

p. 118, lines 15-19. The meaning is not clear here. Is the point simply that more and different adult accounting locations would be required to assess performance if the objective is to achieve a particular distribution of adult spawners in the basin rather than just a total number of spawners?

p. 118, lines 37-38. If the SAR target includes jacks but most analyses exclude jacks, it would be worth indicating how much those analyses underestimate SAR relative to the target. In other words, indicate the typical difference between SARs including jacks and SARs excluding jacks.

p. 119, line 5. Define or explain “Group TWS”; it does not appear to be defined elsewhere in the report.

p. 119, line 8. What are “the two methods”?

p. 119, line 16-17. This section is awkward, seeming to mix a study, a plan, and a company as subjects coordinating efforts.
For both Snake and Columbia River basin PIT-tag salmonid populations, non-parametric 90% confidence intervals are computed around the estimated annual overall SARs for PIT-tagged populations in both the Snake and Columbia rivers.

The subscript in $t_j$ should be fixed.

Define $S_j$.

Indicate (at least conceptually) how $C_0$ will be estimated.

"...fewer parameters (than was the case before 2006) need to be estimated during intermediate steps before arriving at the final annual overall annual SAR estimate as occurred in pre-2006 years.

"In these cases," meaning under the new approach?

"in this group" meaning group T?

Run on sentence is difficult to understand.

"...run reconstruction (RR) of..."

"...we continued extended from 1998-2004 to 1994-2008 the comparison of SARs based on PIT tags and run reconstruction (RR hereafter), 1998-2004, examined SAR methodologies and..."

"...only a single migration year (1999)...". What about 2008? Seems inconsistent with Table 6.1 and Figure 6.2.

"Catherine Creek hatchery spring Chinook..."

"...trends in the overall SARs (LGR-GRA) of wild and hatchery Snake River Chinook groups were similar and highly correlated..."

"...correlated (0.70) during the 1964-2007 period when aligned by smolt migration year."

Is the implication here that the final estimates based on complete returns of B-run stocks may yet exceed the 2% target?

"No PIT-tag SARs have been compiled for hatchery steelhead populations in the mid-Columbia region.” It is not clear whether the effort has not yet been made or whether the requisite data unavailable.
p. 129, line 8. “...however, with creating greater uncertainty in the SAR estimates compared to here than in the Snake River.”

p. 132, line 32. “...inflating average SARs for cohorts that experienced low survival...” Errors in age composition would presumably produce the opposite bias in cohorts with good survival.

p. 133, lines 28-30. Unclear why the high correlation makes continuing the time series important. Is it because hatchery fish could be used as a proxy for wild fish in years when wild fish are too scarce to obtain reliable SAR estimates?

p. 133, lines 31-37. Consider revising to avoid awkward sentence construction.

p. 134, line 23. “…with IDFG run reconstruction SARs for the period 1996-2004, with both time series indicating and SARs from both time series were well short of…”

p. 134, line 37. “…implementing an independent basin-wide independent study of PIT-tag bias study in an effort to evaluate and test…”

p. 135, line 9. “…ocean can suggest which could reveal factors will be that are most critical to survival, and thus provide data as to which actions taken inland will could provide…” Also, regarding uncertainty in meaning, see comment for page 117, lines 25-26.

p. 135, lines 16-22. Consider revising the awkward start of the paragraph. Perhaps it would be better to say: “Additional comparisons of PIT-tag data within seasons suggest that shared environmental factors are influencing mortality rates of both [indicate which species...original sentence says only the two species]. Mortality rates in both species were positively correlated (1) during freshwater outmigration…”

p. 135, line 26. “…a promising line direction of inquiry for upcoming CSS study direction. We plan to explore evaluating evaluate the correlation for of SARs among the regions.”

p. 161, line 7. Replace “we” with “when”.

p. 180 and following. Footnotes for migration year 2008 on Tables B.1 through B.5 should refer to “B” not “A”.