Independent Scientific Advisory Board

Review of NOAA Fisheries' *Life-Cycle Models of Salmonid Populations in the Interior Columbia River Basin (June 28, 2013 draft)*

Appendix - Answers to Questions and Editorial Comments

ISAB 2013-5A, October 18, 2013

Contents

Appendix - Answers to Questions and Editorial Comments

The ISAB's review is provided in two parts: a primary report summarizing findings on each section of the Life-Cycle Model (LCM) document and this appendix, which contains technical and editorial comments. The comments in the Appendix are directed specifically to the modeling teams to help them improve the next model iterations.

Answers to the ISAB review questions are provided below for each chapter and section of NOAA's draft LCM document. Rather than repeat the full questions for each section, the questions are summarized with a few key words, as given in the parentheses below following each question.

1) Are the goals and objectives of the model(s) clearly stated? **(Clarity of model goals)**
2) Are the specific approaches and methods scientifically sound and clearly written? Are there any significant conceptual flaws? **(Soundness of methods and conceptual approach)**
3) Do the models make appropriate use of the data? Are the data sufficient to build and make effective use of the model? If not, what types of data need to be collected to answer the question and develop the model? **(Data use, availability, and gaps)**
4) Is the report clearly written? Are the methods described in sufficient detail for a reader to understand and replicate what was done? Are assumptions and uncertainties about the analyses clearly described? For example, do the authors identify the strengths and weaknesses of the model, and accuracy and precision of model output? **(Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision)**
5) Is the level of complexity of the models appropriate? Does the model output, characterized by metrics on initial conditions, population performance, and population dynamics, allow comparisons among populations and across scenarios within populations? Did they conduct appropriate sensitivity analyses? **(Model complexity, usefulness for comparisons, and sensitivity analyses)**
6) Is the role of the model in adaptive management clearly identified? **(Adaptive management)**

Chapter 1. Introduction

1) Clarity of model goals

This chapter provides a clear and succinct overview of goals and objectives for the life-cycle modeling approach being developed as a common theme for the chapters that follow. However, more context is needed for why this effort has been initiated. How will it answer questions about the effects of habitat restoration, supplementation, flow regime changes, and such? How will the results be used to improve management and/or delist ESA-listed species? The AMIP objectives and the time frame for the next BiOp should be explained. Will the final models encompass the entire life cycle (i.e., including each life stage) and will models represent entire ESUs rather than specific watersheds? What is the timeline for finishing the modeling exercise?

The approach to use this introductory chapter to set up the modeling framework and overall purpose of the modeling efforts is good. While there are many parts of the section that provide excellent information, the section needs some work to make it accessible to those not familiar with fish life-cycle modeling, including definitions of terminology (such as life-cycle stages, model approaches, parameters, etc.), more background on life-cycle modeling in general, as well as previous life-cycle modeling exercises for the Columbia River ecosystem. It also would be helpful to note that the introduction presents the general model setup used in subsequent models and that it lay out the terminology and symbols to be used throughout the report.

It also should be explained that individual modeling efforts focus on the impacts of factors when the salmon are within the Columbia Basin while also (eventually) incorporating variability in survival encountered during marine stages.

2) Soundness of methods and conceptual approach

This is a standard matrix model with stochastic components. Zabel et al. (2006) added a climate forcing function on ocean survival. The authors clearly explain the conceptual history of the current approaches and give practical reasons for developing more complex life-cycle models that are stochastic. The models also include multiple populations, and have the capability to represent density dependent effects and interactions with hatchery fish, and to investigate the population-level impacts of climate change and management actions that improve habitat. The effort to coordinate a systematic approach to modeling these issues, and to develop a common set of metrics for comparison across populations and studies, seems very worthwhile.

Further, the models may be too narrowly focused. One major concern is that the central tenets of the models are based on historical perceptions and information (e.g., flow, temperature, ocean conditions driven by El Nino and PDO). The Columbia Basin has entered a new era where chemical applications, competition and predation by invasive (and sometimes native) species – novel and hybrid communities, ocean acidification, winter icing conditions, and land use are dominant influences on population dynamics. These complexities are not accounted for in the models.

Furthermore, there are quite a few jacks and, at times, precocious males which are currently ignored in the models. Will they eventually be explicitly accounted for in the models, or will they be implicitly accounted for though the “mortality” parameter? There are also a large number of predators in addition to birds, such as non-native fishes. Again are there plans to model these explicitly?

While the modeling approach is standard, and standard terms for fisheries are used, readers who are less familiar with these types of models or who are outside of fisheries may have trouble with the terminology and may not be aware of model assumptions. Here are some suggestions to make the methods clearer:

- Include a glossary of terms, and a clear indication of what the authors are defining as age 1 versus age 2, etc. This terminology should be consistent throughout the report, with exceptions to the conventions laid out in the Introduction and noted in appropriate sections.
- Include adequate references for those not familiar with the approaches. For example, the reference for Leslie (1945) is not included.
- Include a list of abbreviations. For example, ESU may be a term that those new to Columbia River Basin issues may not be familiar with.
- Define all terms in the equations. For example, on page 3, the 4th and 5th equations shown have terms (b_3, b_4) that have not been defined.
- Number all equations.
- Explain and justify assumptions with enough details to show scientific credibility. For example, on page 4, why are jacks typically ignored (see earlier comment)? Is this a justifiable assumption? On page 5, where did the assumption of the quasi-extinction (QE) threshold being 50 spawners per year over any 4-year period come from? How sensitive are the models to this assumption?

Finally, notation is a bit awkward as the superscript refers to the END of a time period, with 0 for a cohort referring to spawning. For example, F_a (fertility) is defined as the number of "age-1" individuals produced by spawners a year after the eggs are laid in the fall. However, in most fisheries literature, fish at the end of their first summer/fall of life are labeled age-0. Similarly, $s_2(t)$ is the survival rate of fish that are between 1 and 2 years old in calendar year *t* rather than survival of fish between 2 and 3 years old. The number of fish in their 4th and 5th year of life (i.e., in the ocean) are modeled as the number of fish that do not breed times their survival in the ocean (s_o). However, this latter parameter is not defined, and the "o" looks like a zero, which leads to confusion.

3) Data use, availability, and gaps

This chapter is a conceptual framework, describing how data are used at a very high level. An appendix describing data sources and data quality would be useful. Furthermore, not all data are of equal quality – a brief discussion of which data are most uncertain and which are crucial for fitting the model is needed.

One of the key uncertainties not discussed is how the model will address correlated random effects that may affect several life states simultaneously. For example, low juvenile survival is often associated with low ocean survival because of a longer-term climate effect. In the model, does each cohort’s cycle act independently of other cohorts’ cycles?

In some cases, it does not appear that all existing information has been thoroughly reviewed or explored. For example, Ted Bjornn and his students did considerable work on survival and stream capacity in places like the Lemhi (e.g., Bjornn 1971), and NMFS biologists have done work to link mainstem Lemhi River flows and survival in that same system (e.g., Arthaud et al. 2010). None of that work is referenced in the relevant study. The lack of credible data to parameterize models and understand the critical uncertainties has been a known issue in the Basin for some time (ISAB 2001-1).

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

This chapter provides a clear and concise overview of the general modeling approach, including major assumptions, uncertainties, strengths, and weaknesses. Nevertheless, more detail would be required for readers to replicate particular analyses; presumably that level of detail is provided in chapters describing specific models. This chapter is intended to describe the “forest” not the “trees,” and it does that well.

While the overview is sufficient to see how the various models will be combined, there was no discussion of model tuning. For example, how are existing data on survival rate to Bonneville of outgoing smolts and ocean survival back to Bonneville used to "tune" the model parameters? What happens if the parameter values for an existing a cohort are used to “forecast” the future? Did the models fit well?

One key assumption is that the modeled life cycle, along with initial abundance, largely determines the probability of quasi-extinction for a population under a specified scenario. Is this a reasonable assumption? Can it be tested using the model against actual data? A Table listing basic assumptions and their possible effects on model outcomes would be worthwhile in the Introduction.

Likewise, a critical evaluation of strengths and weaknesses of the model would help readers. What are possible alternative approaches and what are the strengths and weaknesses of this chosen approach relative to alternatives?

5) Model complexity, usefulness for comparisons, and sensitivity analyses

Life stages in Figure 1 are an appropriate partitioning of the life cycle for an initial exploratory analysis. The authors give practical reasons for developing more complex life-cycle models that are stochastic, include multiple populations, have the capability to represent density dependent effects and interactions with hatchery fish and investigate population-level impacts of climate change and management actions to improve habitat.

The level of complexity of specific models varies depending on the questions being addressed and the availability of information about functional relationships. For example, will the Beverton-Holt relationships used in Figure 1 include the effects of covariates on the productivity and capacity parameters?

The model outputs need to reflect reality and be reasonably predictive. The challenge is to balance complexities associated with life-cycle dynamics, environmental influences on those dynamics, and ease of use and understanding by managers. For example, it is unclear how the risk of quasi-extinction (QE) is extracted from Figure 4. Perhaps a probability scale is missing on the right axis? While Figure 3 and 4 may be illustrative, are not the chances of going 4 years in a row with a mean of 50 or less spawners in a steady state model higher during 100 years than during 25 years? In 100 years there are more chances of "odd events happening." For example, when flipping a coin the chances of getting at least one string with 10 heads in a row is higher with 100 flips than with 25 flips. Perhaps the 25-year QE is just function of burn-in for the model while it settles down? How the model incorporates variability in age at maturity is also important to extinction risk.

There is insufficient discussion of the reasonableness of the model results, e.g., do the results make sense based on monitoring within the Basin? Nor are there any sensitivity analyses or even an indication if they eventually will be considered.

6) Adaptive management

This is not explicitly addressed, but this is essentially the topic of the third paragraph (page 1). The first goal is to model how habitat actions might impact salmon viability. Secondly, the model will be used by managers to see which parts of the life cycle provide the most “leverage” to change total production, i.e. where do restoration expenditures for various life stages improve bottle necks give the biggest return?

The Introduction provides an excellent opportunity to explore the role of these models in adaptive management. Therefore, it would be useful to expand the third paragraph (page 1), and to include more detail about the AMIP. Figure 7 could be enhanced to show how specific models address adaptive management more explicitly.

At the same time, what are the dangers of making changes to management actions based on model outputs that are not accurate? An overview of the possible damages and magnitudes of effects that could result from using faulty model information also should be included.

Specific comments

Page 1, First Sentence: It would be helpful for non-modelers to know and better understand how life-cycle modeling has become “an invaluable tool” for managing at-risk populations.

Page 5: Ocean survival does not appear to be a function of time; i.e., it is simply s_0 rather than s_0(t). So it appears that ocean survival is assumed to be constant. But page 5 says that ocean survival is a function of ocean indicators, which seems to imply that some modeling has been done.

Page 7, Figure 3: Please describe what 5 time series are shown in the figure.

Page 7, Figure 4: It is unclear what the two green lines mean. For these, is the Y axis somehow used to determine the probability of quasi-extinction? If the figure is simply depicting 25 and 100 years as the periods over which QE is calculated, then it can be deleted because the text is sufficient.

Page 8, bottom: Is this common, to model recruitment and then fit an S-R relationship from the modeled data, rather than actual data? Is the purpose of this model to show the uncertainty that the model includes or creates, given stochasticity and density-dependence, under a certain set of parameters?

Page 9: Figures 5 and 6 could be clearer if the dots were grey so that the lines would show up. The caption for Figure 6 indicates a reference to Figure 1 which should probably be Figure 5.

Page 10, top: “The goal of freshwater mitigation actions is to change the state of freshwater ecosystems in such a way to improve the conditions for adult spawning.” Isn’t the goal to also increase rearing habitat for juveniles as well as water quality and environmental conditions for other organisms? This goal is unclear.

Page 10: Specifically, what example analyses are conducted to develop fish-habitat relationships? The last sentence of this paragraph is too vague to know what is being done.

Page 11, top: Section 1.5 needs to be developed more fully, so the reader can be adequately introduced to these topics, which make up a substantial portion of the report.

Chapter 2. Examples of Freshwater Habitat Relations in Life-cycle Models

2.1: Grande Ronde spring Chinook population models

1) Clarity of model goals

The Introduction clearly states that the primary focus is to update and expand the Interior Columbia River Technical Recovery Team's (ICTRT 2007) matrix model by developing more detailed models specific to each of four populations (Catherine Creek, and the Upper Grande Ronde, Minam and Lostine rivers). The ultimate objective is to use the more detailed models to assess natural sustainability and potential for restoration. The Introduction also lays out the organization: the first (and largest) section describes the development of stage-specific functional relationships for juvenile survival in each population; the second section describes modifications to the previous ICTRT model; and the third section describes "*ongoing efforts to further develop the models and associated analyses required to translate independently derived estimates of tributary habitat change into model inputs.*" However, the last section only briefly describes preliminary and proposed efforts to incorporate functional relationships to assess management actions involving habitat restoration and hatchery supplementation. In this sense, the Introduction oversells the current application of this work in that the model is limited in terms of assessing the potential effectiveness of habitat restoration.

2) Soundness of methods and conceptual approach

Development of stage-specific functional relationships and refinements to the ICTRT model appear to be methodical, well-reasoned, and conceptually sound. Probably the most important steps forward are the empirical analysis that reveals strong density dependence in freshwater rearing, especially for the Upper Grande Ronde (UGR) population, and the model refinement to allow simulations that examine changes in juvenile survival (perhaps, as one scenario, reflecting the quality of habitat) and carrying capacity (reflecting the quantity of suitable habitat). The ability to model productivity and carrying capacity separately is a substantial improvement over procedures currently used in the Expert Panel process where a linear relationship is assumed between habitat condition (expressed as a percentage of optimal condition) and salmon survival. However, more analysis could have been done, and is needed, to determine reasons for the strikingly greater density dependent growth and survival of parr in UGR. The remaining comments under question 2 all relate to this concern.

The statistical analysis of various factors affecting parr per spawner seems incomplete (e.g., Figure 3). Shouldn't the various factors be considered jointly to explore interactions? Have confounding time series effects (trends) been ruled out? It remains unclear how the relationships shown in Figure 3 should be interpreted.

In Figure 5, the solid line seems to have been improperly fitted to the Minam River data. Also, the explanation of the dotted and dashed lines seems incorrect in the caption. What is the dashed red line in the upper panels?

The text about the ratio of early to late migrants (page 15, second paragraph) is brief and confusing, and there are no figures to clarify the confusion. Yet differences in this ratio seem likely to reflect, and may be key to discovering mechanisms affecting parr growth and survival. The last sentence of this paragraph seems especially surprising and warrants explanation: if the measured ratio of early to late migrants is an artifact of the placement of traps, then doesn't this uncertainty about representativeness of samples cast doubt on the entire juvenile enumeration effort and all survival analyses based on these enumerations?

Also on page 15, it is stated that differences in average length of fall and spring migrants could be explained by differential mortality of smaller fish or growth in winter and early spring. Could growth increments of PIT-tagged individuals (i.e., size at release versus recapture) be examined to decide between these alternatives?

In general, the statistical analysis of relationships involving length could be improved. For example, measures of statistical significance are lacking in regressions in Figure 8; are the indicated slopes significant (i.e., credible) in all plots? It would be worth comparing and presenting the length distributions for migrants by population; the scatterplots in Figure 9 suggest that UGR migrants are smaller on average, which would be consistent with density dependent factors affecting their overall survival to Lower Granite Dam.

Neither the main text nor Appendix C mentions whether the AQI procedure to standardize for habitat area takes into account water temperature, limiting flow rate, or other factors beyond channel form, flow characteristics, and slope. It seems not, in which case the amount of ***suitable*** habitat in UGR might have been overestimated relative to other streams. Overestimating the amount of suitable habitat in UGR during the AQI standardization step could also explain the steeper, extrapolated negative slope in Figure 10, giving the appearance of greater density dependent survival in UGR. In other words, reducing the estimate of suitable habitat in UGR (relative to other streams) would increase calculated densities, moving them proportionately to the right in the plot, and decreasing the slope of the relationship. Note that Crozier and Zabel (chapter 2.4 in this report) found that summer stream temperature and flow exerted a primarily density-dependent effect on parr-to-smolt survival; that is, they detected a greater statistical effect of summer stream flow and temperature on the Beverton-Holt parameter for carrying capacity than on the parameter for intrinsic productivity.

Alternatively, might the explanation for the greater density dependence in UGR be that recent densities of hatchery fish spawning naturally (pHOS) are higher in UGR than in other streams?

Such an explanation would be consistent with Buhle et al.'s (Chapter 4 in this report) conclusions about the lower Rmax of hatchery-origin fish.

3) Data use, availability, and gaps

This study is a major contribution in synthesizing and documenting field data collected with many years of effort. In general, the study appears to have used these data effectively, although in some cases, it is difficult to judge due to lack of clarity in the presentation.

The authors should clarify how they dealt with gaps in annual flow records (page 6, paragraph 1). According to the authors, "*there were gaps (one to three years duration) in the annual flow records for the other three populations*." Did they estimate flows for stations where data were not available in a given year from a surrogate station based on relationships among stations in past years? If so, was the site used a good surrogate for estimating flows in the unmeasured sites based on past years? Or was a multivariate approach used to estimate missing flow data from past records at the missing station and other adjacent stations?

What is known about the accuracy and precision of the redd counts used to estimate spawner abundance?

The authors do not discuss the accuracy of estimating juvenile emigrations from weekly estimates of capture efficiency for rotary screw traps. Efficiency estimates are a weakness in many screw trap studies when the expansion is large, because errors in the efficiency estimate can lead to large errors in the total population estimates. Consequently, efficiency estimates should be presented as well as the degree of expansion used. It would also be good to clarify the adequacy of weekly efficiency estimates (e.g., how quickly stream flow changes) and to indicate how errors in those estimates might affect the assessment. In particular, the downstream movements of smolts may be different for fish gradually drifting downstream to overwintering sites compared with those actively emigrating as smolts. Recoveries of marked fish used to estimate the total population of migratory fish may be biased differently depending upon their stage and migratory behavior.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

Overall, this document is a good first draft in terms of technical content, but it is difficult to read and comprehend in many places. It would benefit from editing for clarity and perhaps some reorganization (see Specific Comments below).

The methods are couched at an appropriate level of detail, but they are not always clearly conveyed. Assumptions are usually mentioned in the text but often lack justification. Scientific literature is cited infrequently.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

The models are just complex enough to estimate the effects of particular interest, yet remain simple enough that data are available to include and compare multiple populations with contrasting histories of supplementation. The five habitat scenarios in which the survival rate of parr and the amount of rearing habitat are varied selectively and systematically (increased and decreased by 10%), constitute a simple sensitivity analysis. Additional complexity may be needed to investigate the impact of hatchery supplementation programs, as listed under Next Steps; the modeling approach and findings in Chapter 4 may be useful in this effort.

6) Adaptive management

Adaptive management is not discussed explicitly. However, the ultimate objectives for updating and expanding the model to test hypotheses about the impacts of habitat restoration and hatchery supplementation and to determine the extent of habitat improvements necessary for recovery are consistent with the adaptive management cycle in the Fish and Wildlife Program. We encourage the authors to recast some of their discussion with this in mind.

Specific Comments

The Introduction should provide some more perspective on the theme and context for what is to follow, i.e., the motivation for the work.

Methods are sometimes misplaced. For example, the following sentences occur in the beginning of the Results (pages 9 and 10): "*We standardized each of the four data series to spawner and summer parr per 10,000 m2 of pool habitat using estimates from the ODFW Aquatic Inventory (AQI) surveys (Table 1). For each population, we expressed the results as an AQI index of pool equivalent habitat by weighing the category habitat subtotals by the relative density index for each category (Table 1:2.1 Grande Ronde Chinook Population Models10 pool, runs and fastwater). We used the resulting population totals to standardize spawner and parr densities to a common unit of habitat. For Catherine Creek, we estimated an additional expansion factor to account for the use of habitat below the weir site for spawning and early rearing.... We compared summer parr per spawner ratios (standardized to 10,000 m2 AQI habitat) against the flow and temperature indices and against parent spawning densities (Figure 3).*" Reorganization could improve readability and flow of the paper in other instances too (e.g., page 12 and page 15, paragraph 4).

Figure 1 - Trap 5 (Elgin) is missing.

Page 7 (and 27) - The procedure for standardizing density is not clear. Undefined computer variable names (only some of which are self-explanatory) are used instead of proper mathematical notation. In this document, mathematical procedures are described in words instead of equations, which can sometimes enhance readability, but at the great cost of imprecision such that readers are unlikely to be able to repeat any particular analysis.

Pages 7-9 - The verbal description of parr-to-spring migrant and smolt migrant-to-Lower Granite Dam survival calculations is confusing. Abundance and survival terms are sometimes mixed within sentences, making the logic difficult to follow. For clarity, it would be worth using equations and defining a simple specific vocabulary to refer explicitly to the various stages of survival that are discussed, and to use this vocabulary consistently throughout the report.

Page 10 - In the section on “Flow and Temperature Indices,” the last sentence (regarding density dependence) is out of place; it should be moved to the previous section. Also, the text before states that there is no significant ***trend*** in parr per spawner, but no trend data are shown, which seems like an oversight considering the level of detail presented in other figures.

The caption to Figure 8 contains an error regarding the right-hand panel.

Figure 12 is informative, but the labeling of units could be improved. In the right-hand panels, are spring presmolts and Lower Granite smolts in units of AQI density or total production for the population? An equation or more explicit text (on page 27) is needed to show precisely how summer parr AQI density units, survival functions, and rearing areas are translated into total smolt production.

The five habitat scenarios should be explained in the caption to Table 7 and in the text preceding it. (They are listed later in the caption to Figure 15, but it took time to figure them out).

The second and third sentences of the Discussion are confusing. What is the “survival relationship”?

2.2: ISEMP Watershed Model for spring/summer Chinook salmon and steelhead in the Salmon River Subbasin

1) Clarity of model goals

The objectives of the current analysis should be stated more explicitly in the Introduction. However, it is clear that the current analysis is intended to simulate habitat restoration scenarios and to predict corresponding changes in adult Chinook abundance. The Introduction provides a useful explanation for the need to develop life-cycle models, and how the intent of this report is to describe improvements to an earlier version of the ISEMP Watershed Model. The ultimate goal is to apply a generalized version of the Watershed Model in other subbasins and to guide monitoring design.

2) Soundness of methods and conceptual approach

The approaches and methodology used to adapt the models specifically to the Lemhi are not adequately described. Although the modeling is obviously an extension of the general work outlined in Chapter 1, the specific modifications for the Lemhi, the nature of the assumptions, and the logic for parameterizing the models and scenarios could be improved substantially. Significant details are missing, and the terminology is often confusing due to inadequate definition, apparent errors, and inconsistent usage. To improve clarity, equations should be written as mathematical expressions (as in equations 1 and 2) instead of computer script or line style (as in equations 6 and 7).

How do the authors know the models are working as expected? Were results checked against any measured data?

3) Data use, availability, and gaps

We are unable to judge whether the models make appropriate use of existing data. The current ISEMP project is cited and presumably provides the bulk of the available data, but the use of those data is not clearly explained. A body of other work in the Lemhi also exists (e.g., by Bjornn 1971, Arthaud et al. 2010), but it isn't cited and presumably isn't used. Overall, it is unclear what data are used, how they are used (i.e., to develop relationships or as parameters for the model), and whether the model outputs have been checked against any measured data. In sum, the report does not identify what we know, what we don't know, and what the most important gaps are.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

The most obvious deficiencies in the current draft are lack of clarity in the writing and inadequate detail about assumptions and how data are used. Many assumptions are not obvious, and the logic is not always clear. A reader could not replicate this study from the description provided. In the end, it is not clear which results are reliable and which are speculation, or how this Watershed Model is connected to NOAA's overall initiative to develop life-cycle models in the Columbia River Basin. We suggest the authors to examine Section 2.3 of the LCM report as a good example of how to describe methods.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

Although many aspects of the current model application are difficult to understand, the level of complexity of the model itself does seem appropriate. The problems are chiefly with the application of the model, the detail and justification of the parameterization, and the lack of well defined uncertainty or sensitivity analysis to guide further work.

6) Adaptive management

The Introduction states that the ultimate goal is to apply a generalized version of the Watershed Model in other subbasins and to guide monitoring design, but its role in adaptive management is not discussed. Models like this can and should provide an important context for further monitoring and research, and when more fully developed and tested, may help to ascertain the extent of habitat restoration needed to restore population viability.

Specific Comments

Title - The paper is really about Chinook so the inclusion of steelhead in the title seems unnecessary. Subsequent discussion of model refinements for steelhead also might be dropped or abbreviated to simplify presentation and improve clarity.

Introduction - Simulating habitat effects on survival is not supported later in the paper; it is not clear how habitat changes were simulated other than through an arbitrary adjustment of the carrying capacity (or productivity?) parameter. To be instructive, the specific assumptions of how changes in habitat influence survival need to be explicit and should be communicated in the Introduction to aid the reader later on.

No application of source-sink or recolonization dynamics is discussed later in the report, so this point could be dropped or clarified to prevent confusion.

Page 2-3 (Beverton-Holt function) - This section needs some clarification. The section heading (or perhaps additional subsections) should be set up to fit the discussion of successive life stages that follows. More detail would be helpful here – what life stages are being modeled as density dependent with B-H functions? A schematic of the model showing the different life stages and departure from the general model would help. Detail is also needed with the logic and assumptions; for example, how are productivity and spatial structure related?

How are the scalars $E_{i,q}$ (that reduce maximum productivity at each site) known? How are the maximum survivals and densities known? Presumably the $Sr_{k,i,t}$ correspond to the *intrinsic* productivities (i.e., maximum productivities at very low density) in the highest quality habitats, but how are these values known? Similarly, are the $D_{k,j,i,t}$ known empirically or estimated in the model? How is it known that the $c_{k,l,t}$ (calculated from H and D) are maximum values for the site and stage in question?

The report claims that the model is grounded in empirical data, but the empirical basis for the parameters driving density dependence is not shown. The focus seems to be on habitat area or capacity, but how do these variables determine survival (or factors that might influence survival) at each of the life stages that may be relevant. What is the logic (i.e., functional relationship) for linking habitat change to survival and how can one estimate the changes?

The description here is a departure from the general model discussion presented in Chapter 1, apparently because more detailed information is available for the Lemhi. Presumably, the details are available in the QCI report, which is not readily available from the linked citation. Can the authors link to the report directly and does it fill in the gaps? Is the QCI report destined for publication? It seems like a large amount of work is being extended from a potentially uncertain foundation, or at least one that has not been peer-reviewed.

Some terms are defined inconsistently (and it seems incorrectly). For example, $H_{k,j,t}$ is defined as "the fraction of habitat j at time t in site k" in the equations but as "stream specific habitat characteristics" in the text; $D_{k,j,i,t}$ is defined as "the area (m^2) of habitat j at site k during life stage I at time t." Instead, it seems instead, from equation 4, that $H_{k,j,t}$ must have units of area (as in $A_{k,t}$), and from equation 3, that the variable $D_{k,j,i,t}$ refers to the maximum density of life history stage i within $H_{k,j,t}$. In any case, the description is confusing. "Spawner productivity" is used to indicate fecundity, but elsewhere, productivity is defined as smolts or parr per spawner.

Pages 3-5 (Modeling parr to presmolts) - This discussion seems unnecessary since it does not specifically inform findings with regard to Chinook salmon.

The sentence after equation (8) is confusing: *"we assume that survival from one pre-smolt stage to the next is proportional to the number of fish attempting to enter each pre-smolt age and the survival probability of the candidate group transitioning to pre-smolt."* It seems implausible to assume that ***survival*** is proportional to ***number*** of fish.

The phrase after equation 10 seems incomplete: "*but must be scaled such that the total equivalent first year pre-smolts.*" In equation (11), a subscript seems to be missing between the two commas of $N5_{k,,t+1}$

The values being assumed should be documented and justified. It would be helpful to provide a table indicating the source of values and citing other reports or studies as appropriate. It also would be helpful to indicate which parameters are "fitting" parameters and which are to be manipulated in the Monte Carlo simulations. See Table 3 on page 11 of Chapter 2.3 for a good example of such a table.

Page 6 (Cross-Site Migration) - This paragraph should clarify whether the intention is to develop a metapopulation model or a spatially explicit population model with fish moving across habitats through different life stages. If the former, it isn't clear how all the component populations are tracked and modeled, individually or as a composite, through all the relevant survival functions. If the latter, it is not clear how capacities and survivals are established for all the different habitats. How is the fraction of the population at a given life stage determined? Movement probabilities are later presented (in Table 7), but it is not clear whether they are derived from this cross-site migration option, and the use of different terms is confusing (e.g., migration probabilities on page 6, juvenile migration rates in Table 2, movement probabilities in Table 7).

Page 7 (Stochasticity) - The issues of uncertainty and variability are important, but there is no detail here (or later) about what information is used, or how it is used, to consider the implications of uncertainty and variability.

The last line in the paragraph following Equation 20 says “Varying μ_t and λ” but should μ_t be μ_{target}?

In the last paragraph before “Data Requirements”: “discreet” should be “discrete”; the sentence “*For both temporal trends and step functions, within-run quantile probabilities carry over from the initial mean values to target mean values, and all variability components and correlations are preserved as temporal trends or step functions are imposed*” becomes unintelligible, starting “*within-run quantile probabilities*”; “what is “it” in “*it would continue to use the 78th percentile*”; it’s not clear how quantiles are “consistent” within Monte Carlo simulation runs.

Figure 1 - delete “that” or insert “was” in the caption.

Page 9 (Adult and population monitoring) - More detail on the resolution and structure of the model is needed in this section. For example, the sentence to “*These various survey types allow estimates of sex ratio, age structure, abundance, survival, and habitat associations at egg to smolt life stages, at broad (population) and fine (individual treatment and reference sites) scales*” is vague and does not explain how life stages are being modeled. Taking care to clarify meanings precisely (rather than generalizing) in sections like this might help.

Pages 9-10 (Modeling habitat effects, last sentence) - We appreciate the value of modeling how a change in habitat leads to a change in survival, but we don’t see any explanation of the assumptions or functional relationships that would allow this modeling to be done in an objective way. We can guess how capacity might be adjusted in response to improving access to a stream or adding new habitat, but we do not see any explanation for modeling changes in survival. More detailed discussion of the assumptions and relationships is required here or in the application below.

Page 11 (Background, first paragraph, last sentence) - See preceding comment regarding anticipated changes in survival. What the Action Agencies anticipate is largely speculation, so it would be useful to provide more detail about how habitat actions are expected to improve survival. It’s clear that reconnection should increase capacity, but it isn't clear how increased capacity will change survival by the magnitude indicated. Are the increases of 4-7% additive or proportional? Also didn't earlier work by NMFS in Idaho indicate that flow in the mainstem

Lemhi River had a major influence on survival (see Arthaud et al. 2010)? There is no reference to that conclusion or the ability to consider it in the model. This work seems focused on a handful of things that can be estimated, but doesn't consider (or discount) all the other things that might also be important.

Page 11-12 (Scenario descriptions) - There is a lack of detail. The assumptions are not apparent and the logic is not clear, even though this section includes the guts of the application for linking habitat restoration to population response. How do the assumptions regarding habitat change influence the selection of variables and how are they justified? What is known well and what is just guess work? For example, how does the productivity scalar reflect competition? What changes and by how much? What information is available to support the assumption? If all the early life history is represented as a series of B-H functions, how are the survivals adjusted (or not) in the different scenarios outlined in Table 3? The first sentence on top of page 12 states that Restoration Scenarios 1 and 2 differ only in ***carrying capacity***, but the caption to Table 2 says they differ in ***productivity*** (50% versus 100% of Hayden Creek's value). The difference is critical to the logic of the analysis, but it is not clear which statement is correct.

Table 2 - Are "juvenile migration rates" the same as migration probabilities (mentioned on page 6) refers to "movement probabilities" (in Table 7); do these rates have units or are they probabilities? (Consistency in terminology is important with anything as complex as this subject matter.) Some more detail to explain what is in the table would help. What is the "target value"? What is "tributary productivity" and how does it relate to other "productivities," i.e., smolts per adult? If productivity actually means survival, then for what stages? See our earlier comments regarding details on the B-H and life history stages. Add parameter symbols to the tables to clarify how these values relate to the equations.

Table 3 - Given the earlier discussion about model structure, we expected to see measures of the uncertainties and variability associated with estimates, as well as information on how values might change under alternative scenarios and assumptions about competition or habitat quality and capacity. Many site-specific survival rates in this table are identical, suggesting that they are assumed (or perhaps generalized from one site to another). The claim that the model is grounded in empirical data is weakened to the extent that this is true.

Table 7 - Should the probabilities sum to 1 within rows (some do not)? What data are used to derive these probabilities, and how are they calculated?

Pages 24-25 (Discussion) - The last sentence of the first paragraph on page 25 seems contradictory. How can the model be capable of detecting change without error structure and an analysis of uncertainty? Perhaps “capable of detecting” should be merely “capable of simulating.” Some discussion of this distinction seems important.

In the last bullet point, the use of the term “productivity” is again confusing; previously and subsequently, productivity is defined as “smolts per female.” The B-H functions seem to be derived from capacity and survival information, but it is still not clear how these assumptions are implemented or whether they make sense in terms of the way productivity is defined. How can productivity be defined per area? That seems to imply something about capacity.

2.3 Upper Columbia River spring Chinook salmon

1) Clarity of model goals

The goals and objectives are not clearly stated. While sentences explaining intentions are stated throughout of the report, the clearest indication of a goal is on page 4: “*One of the goals of developing this model is to characterize the impact of habitat actions on survival and subsequent impacts on population dynamics.*” Presumably this is a longer term objective, as the issue is not addressed in this report. The Introduction would be improved by stating the issues to be examined; how the model will be linked to habitat, either natural or restored; the ultimate reasons for developing the model; and how the model will be used to improve decision making, habitat restoration efforts, or assist with the delisting of ESA-listed fishes.

The authors provide some useful background about hatchery programs, and indicate that future applications of the model will assess population responses to habitat actions, hydropower system operation, pinniped predation, and climate change. However, they do not state what the current models are being developed to address.

2) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

Overall, the approaches are described clearly and seem sound scientifically. The biggest deficiency is that the work to date has not clarified fish-habitat relationships for Wenatchee populations. Instead of predicting population responses to habitat restoration actions, the simulation scenarios merely illustrate various arbitrarily assumed consequences for survival. It seems that the data or conceptual tools are still inadequate to examine quantitatively how freshwater habitat causes changes in Chinook abundance, recruitment or extinction in the Wenatchee system.

3) Data use, availability, and gaps

The models that are presented (i.e., developed to date) appear to use the requisite data appropriately. However, empirical data are not presented on fish-habitat relationships. Are these data available from ISEMP or CHaMP programs in the Wenatchee and Entiat rivers? We presume that models using these data are still being developed and will be presented in the near future.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

The report is clearly written and logically presented, with good succinct use of supporting equations. The focus would be improved if the objectives (see comments above) and the logic for the basic model setup were described more fully in the Introduction. The Discussion would be improved by reviewing assumptions and uncertainties in the current approach, and by considering the implications of these limitations.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

The level of complexity seems appropriate to the simulations described in this report and should be adequate for those listed as planned. That said, the current model is not very complex because the model does not yet include functional relationships between habitat characteristics and fish survival or density.

A complete and methodical sensitivity analysis was not conducted, but the simulation scenarios provide a basic sensitivity analysis in which selected parameters of interest were adjusted separately to determine outcomes for the fish populations.

6) Adaptive management

Adaptive management was not discussed. The Introduction would be improved by stating the ultimate reasons for developing the model, and by considering how the model might be used in the adaptive management cycle to improve decision making for habitat restoration and to assist with the delisting of ESA-listed fishes.

Specific Comments

Introduction - Use of “supplementation from a segregated hatchery” might cause confusion, as in the Columbia River Basin, supplementation often refers to an integrated hatchery program.

How can life-cycle models be expected to capture population responses to relatively small changes in freshwater habitat in the face of large year-to-year and decade-to-decade variability due to other factors? Consider, for example, the magnitude of variability documented by Hilborn et al. (2003) and others for sockeye salmon populations in Alaskan streams – streams that had not received any habitat perturbations or restoration. Also the high variability in local fish responses to habitat restoration (i.e., changes in fish density or abundance) means that large restoration efforts are needed to produce measurable changes in fish abundance at a watershed scale (Roni et al. 2010).

Page 4 - Does the Basin include sections with no hatchery influences that can be used as reference sites?

- The analysis was restricted to spawner data from 1992-2007. Explain why it is a problem that "*1991 and earlier years reported a constant proportion of spawners allocated to each age class*".

- It's not clear why initially (pages 4-7) alternative versions of the Ricker model are used to test for population-specific parameters for intrinsic productivity (a_i) and carrying capacity (b_i), as well as hatchery proportion and year effects, but subsequently (page 12), a Beverton-Holt model is fitted to each population in the simulation model for scenarios. For clarity, the parameters for the two models should be represented by different symbols; also Y_t should be defined. On page 12, the verbal description of how parr abundance was reconstructed from subsequent adult returns is confusing; the equation helped to clarify the text.

Page 6 - To assume that detected differences among populations in productivity (survival at low density) and carrying capacity are habitat dependent begs the original question. These differences could also result from other factors such as disease, predation, hydrology, and so forth. Are there habitat data to justify the assumption? The authors need to be more specific about all types of landscape change used in the modeling effort. More work is needed to identify why the "year" effect was always significant; i.e., what is the associated variable(s) that actually kills fish? Capturing this mortality with the proxy variable "year" may obscure the influence, and thus prevent the discovery, of functionally more important variables.

Page 8 - Are there data to support the statement that "*Subyearlings leaving their natal tributaries for the mainstem Wenatchee River experience a warmer thermal regime and less competition for habitat and food resources than the yearlings that rear in the tributaries, both of which may contribute to the higher growth rates observed for subyearlings compared to yearlings (WDFW, unpublished data).*" The differences in size may have more to do with bioenergetics.

Figure 6 requires more explanation. Are the plotted points the average length for emigrants in each year?

Table 3 is helpful, but raises a few questions. The *b* parameters for the Beverton-Holt function are identical for populations in Nason and White rivers. Does this mean one or both are assumed rather than estimated separately? If so, the authors should explain why. The authors should also explain the derivation of the σ^2_1 and ϕ_1 shown in Table 3.

Page 13 - It's not clear why values for first year ocean survival, s_3, are computed with the equation using input values for water travel time (WTT) and the two upwelling indices (April and May) drawn randomly from recent, bad or historical distributions. Why not just randomly draw values for s_3 from the corresponding time periods?

Page 14 - The third sentence states *"'Recent' referred to the recent period (1980-2006) of environmental time series, 'bad' referred to values drawn from the period 1977-1997, and 'historical' referenced conditions observed over the entire length of the environmental time series, 1946-2006."* Is this correct? There is significant overlap between the "bad" period and the "recent" period. Are these periods so different that good/bad labels can be accurately applied?

Page 15 - Why are negative estimates of k-values reported as NA?

Page 15 - in the last sentence, clarity would be improved by adding "Reduced" before "avian predation…"

2.4: Population responses of spring/summer Chinook salmon to projected changes in stream flow and temperature in the Salmon River Basin, Idaho

1) *Clarity of model goals*

The focus of the study is clearly stated in the Introduction: "*In this report, we focus on developing techniques and analyzing the interplay between density, temperature, and flow from spawner to smolt stages. Future work will incorporate additional migration and marine influences. … Here we build upon previous analyses with additional data on the influence of climatic drivers, specifically temperature and steam flow, on parr-to-smolt survival in spring/summer Snake River Chinook salmon …*"

The authors also clearly indicate that the work is preliminary and exploratory. However, it is less clear how this study meshes with the overall goals for the Life-cycle Models report, or the other models described in Chapter 2 that are being developed to incorporate functional relationships between fish survival and changes in habitat. It would be useful to clearly summarize the goals and objectives, stating how the objectives, once attained, will be used to improve management actions.

2) Soundness of methods and conceptual approach

The conceptual approach in this study is innovative, ambitious, and clearly explained. The study achieves its stated objectives of "*developing techniques and analyzing the interplay between density, temperature and flow from spawner to smolt stages.*" However, the modeling is (as yet) too limited in scope to provide realistic results that can guide habitat restoration or other management actions. For this reason, the authors warn that the model results should not be used as predictions for decision making.

The logic of the analyses and simulation modeling appears to be sound, although some of the assumptions can be questioned. The following questions and comments indicate where more justification, explanation, or discussion would be helpful.

No doubt, at one scale, there is a good correlation between tributary and mainstem stream flow, (as explained on page 7), but that scale may not be appropriate for understanding parr-to-smolt survival. Short-interval variations in tributary flow can have important consequences for survival, and these perturbations may not be captured well by using data from the mainstem Salmon River. This important assumption needs to be tested. Furthermore, temperature recording stations do not capture spatial variability in temperatures within a stream reach. One would like to know the temperature environment experienced by the fish rather than the temperature from a set station. Fish are very good at finding comfortable temperatures, if they are available.

Do the estimated differences among streams in direction and magnitude of flow, as well as temperature effects on carrying capacity (c2), make ecological sense (i.e., are there plausible mechanisms) or might they just be statistical artifacts reflecting the confounding effects of other excluded variables? Relying solely on statistical relationships without understanding the mechanisms involved is risky. In-depth exploration of the mechanisms that are assumed or implied, and how they might be validated, is needed.

Anecdotal evidence for earlier migration out of headwater habitat in warm years (which might cause crowding at lower elevation) is mentioned (page 35) as a potential mechanism that could account for the larger climate effect on c2 than p2. This suggestion does seem plausible and warrants more discussion. It might also be worth considering evidence for, and the potential effect of, life history diversity in migration behavior, both within and among populations. As an alternative explanation for the greater temperature effect on c2 than p2, is there evidence that variations in temperature among pools or minor tributaries within identified streams might be sufficient for overall stream or air temperature (at the level measured in this study) to change the relative amount of habitat within a stream that is suitable for rearing? (We have suggested that such a climate-related reduction in suitable habitat might explain why density dependence during freshwater rearing is so much stronger in the Upper Grande Ronde than in other populations of spring Chinook in the Grande Ronde Basin, as described in chapter 2.1 in the LCM report.)

Regarding the statement (page 35, second paragraph) that “*populations heavily influenced by hatchery inputs, which have external drivers of productivity (e.g., Buhle et al. 2013) and potentially very high fish densities,*” it is worth noting that Buhle et al. detected stronger effects of hatchery proportion on carrying capacity than on productivity of naturally spawning Chinook, similar to the findings for climatic variables in this study.

The survival analysis and life-cycle modeling depend heavily on fitting relationships with a limited understanding of the underlying mechanisms. Although differences in results across populations may indicate fundamental differences in mechanisms by stream, it may also indicate substantial uncertainty associated with the limited data. One concern is the timing of the (statistically) most obvious flow and temperature effects on parr survival. What mechanisms could link summer temperatures and flows to parr survival that takes place after the summer period? Certainly growth during summer could be related to subsequent survival, and size is an important predictor. Is that the presumed mechanism for the inclusion of summer temperature in the B-H models? Some further discussion of the indirect nature of the apparent relationships and the assumed mechanisms could help strengthen the discussion. Without a clear notion of the mechanisms, forecasting associated with the climate modeling seems a bit of a stretch.

It is stated (on page 36, second full paragraph) that increased summer temperature also can have a compensating (positive) effect on survival by increasing growth, and hence fish length. If this positive effect is strongest at low fish density when more food is available per fish, as seems likely, then the overall negative effect of temperature on survival would become more evident as density increases, and this relationship could explain why temperature showed up as a significant negative effect on c2 rather than p2.

The sentence (page 36, top of page) “*Some warm streams might have greater capacity for increasing primary productivity, which could raise the optimum temperature for salmon growth.*” raises the possibility that deliberate interventions to modify habitat productivity or capacity might increase the resilience of some populations to accommodate increased temperatures. It might be helpful to comment on the feasibility of this option, or to recommend the research that would be needed to explore it.

It isn’t clear what is meant in stating (on page 36, first full paragraph) that the stronger positive correlation of stream flow with parr-to-smolt survival in the mark-recapture data for PIT-tagged fish, compared with reconstructed data for spawner-to-smolt survival, "*might reflect observation error on our estimates for smolt productivity.*"

3) Data use, availability, and gaps

The models make extensive use of relatively limited data by building on earlier work with apparently different assumptions. More discussion of data limitations and the modeling would be useful, given that some conclusions are based on complex extrapolations from simulated relationships based on statistical correlations rather than knowledge about ecological processes.

More consideration should be given to how errors (or uncertainty) in estimates of fish abundance might affect conclusions that are based on statistical differences in how well complex models fit the (questionable) data. Fish abundance is difficult to estimate precisely, and survival estimates are particularly prone to error because they are a ratio of two estimates of questionable precision. For example, Zabel et al. (2006) say there are not reliable estimates of the number of parr. For this analysis, the number of parr is estimated as a latent variable (using equations 3 and 4) based on spawner redd counts (see our comments in next paragraph), estimated smolt abundance, and measured survival of PIT-tagged parr. Smolt abundance itself is estimated (reconstructed) from redd counts in a chain of assumptions based on estimates of age composition and harvest rate (equation 8), and SAR and prespawning mortality (equation 9). There is a lot of room for propagating error in these procedures. This comment is not intended as a criticism of the innovative modeling that offers an opportunity to infer relationships of interest in these data, but rather, to ensure that the inferences are based on signal rather than noise.

The authors point out that the most important data driving the parameter estimates are population specific spawner counts from the Salmon Population Summaries. We are not aware that there are any spawner counts for these systems - only redd counts, and many of those outside the Middle Fork proper are only index counts. Even Isaak and Thurow's (2006) complete census of redds in the Middle Fork is based on only one survey during the spawning season, and has not been corrected for observer error. Have the authors considered the potential for error in estimates of spawner abundance or the implications of such errors for the conclusions?

What are the implications of the values used for upstream survival and age-specific fecundity? How certain are these values, and are the models sensitive to errors in these values? Also, what are the implications of excluding precocious males from the matrix model (page 13)?

Perhaps the biggest concern is that empirical data on parr density. Experimental support for functional relationships between fish survival and habitat variables (and other components of climate change) are not available in any of these streams to independently test predictions from the statistical models. Further, the authors have been restricted to a limited view of factors affecting survivorship because data to model more realistic relationships (e.g., those involving winter ice, disease, predation, competition from non-natives) are not readily available. Winter icing conditions (about which there are no data) may be a key driver of survival of eggs and parr and, as well, it would be useful to know how fish condition factor or fat content influences survivorship.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

In general the report is well written and the method sections provide important detail. Many of the assumptions and uncertainties are acknowledged and discussed, but the basic limitations of extrapolating from unreliable data remain troubling. It would be appropriate to include a table summarizing the assumptions, and indicating the sensitivity of model outcomes to them. The Discussion includes useful and insightful speculation about possible mechanisms and implications, but this speculation should be clearly labeled as such and should not be mixed with the quantitative conclusions.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

The level of complexity in the models seems appropriate given the restricted scope of this analysis. The models provide interesting comparisons among populations and across scenarios, and provide information on the sensitivity of results to changes in the critical parameters (e.g., flow and temperature). However, it is not clear what level of complexity is required to obtain critical information for better estimating parameters used in the simulations, and for testing their potential influence on the conclusions from this analysis.

The current analysis of climate change is quite restricted in the sense that only stream temperatures and flow are modeled, and only at coarse scales. Consequently, the current analysis may be inadequate to elucidate realistic mechanisms driving survivorship of salmon over the full life cycle. The authors indicate that other information on migration and marine influences will be incorporated in future efforts. Regarding marine influences, climate change affecting the size of the Columbia River plume and acidification of the ocean, with concomitant effects on the marine food web, may have larger impacts than surface temperature.

While the authors acknowledge limitations to the scope of this analysis, we think it should be emphasized that the current analysis lacks explicit consideration of habitat, landscape change, land-water linkages, competitors, chemicals, or ocean acidification. The last four themes will profoundly affect future conditions (Davis et al. 2013; Lucas et al. 2013). This may be the reason why, as the authors state, that "*The more equivocal nature of these results compared with the survival results directly suggest more complicated dynamics at the population level.*"

6) Adaptive management

Adaptive management is not explicitly addressed, perhaps because the focus of the study is really to develop techniques for analyzing the interplay between density, temperature and flow from spawner to smolt stages. The authors make it clear that the results should not be used for management decisions at this point. Nevertheless, at some point, models like this will be developed more generally to guide decision making, and their role in adaptive management should be considered and promoted.

Specific Comments

Page 4 - How large is the hatchery influence in the South Fork Salmon River? Might this have caused a positive correlation with temperature when density was included instead of length (page 10)?

Methods for Part 1

Page 7 - Explain how cumulative growing degree days in the stream are defined in reference to the Crozier et al. (2010) study. Is this definition species-specific? The reference to Achord's team is unclear when first mentioned, because their work is not yet cited.

Page 8 - why use modal length instead of mean or median length?

Page 9 - "equation 11" should be "equation 1." Under Analysis, it is redundant to list predictor variables again (suggest consolidating this text with existing text on pages 7-8).

Methods for Part 2

Page 16 - The Executive Summary refers to a Bayesian approach, but the approach is not clearly defined as such in the Methods. It would be helpful to more clearly explain how model fit was assessed based on the two likelihood components. Some related jargon might be unclear and is worth defining more carefully (e.g., normal hyperdistribution and MCMC chains).

Page 18 - Indicate the source of the information for the upstream survival and age-specific fecundity parameters (i.e., why is the in-river harvest rate 7%, why is 0.7254 an appropriate constant for natural survival, why is 1.26 the number that got multiplied by 5-year old spawners)?

In Figure 3, are the means of the posterior hyperdistributions for the temperature and flow effects much different than zero? Can the significance of this difference be quantified or expressed statistically?

In Table 3, it seems that most of the flow (B2) coefficients are not significantly different from zero; can this be tested and indicated? Also, for consistency and clarity, it is better to use β as in text rather than B; label columns as log(p1), log(c1) and logit(p2), include units for these values as appropriate; and spell out SD (not immediately clear what second table shows).

Page 20 (Ocean conditions) - More explanation is needed about how the starting year in historical series is randomized in the simulations. Why randomize the start time, rather than draw years randomly with replacement (i.e., bootstrap)?

Page 20 (Simulations) - The first sentence appears to be about meteorology, so the second sentence about population dynamics seems to be misplaced. On the previous page, 10 GCMs were mentioned, yet sometimes only 7 GCMs were tested. Some explanation about why all 10 GCMs were not tested would be helpful. Also in Figure 4, it would help (first time) to indicate that the colored lines and abbreviations in the legend represent the GCM scenarios; those unfamiliar with GCM models will find the abbreviations of GCM models rather cryptic (some explanation would be helpful); also add units for temperatures (presumably degrees Celsius). In Figure 5, it would be useful to include horizontal reference lines to facilitate comparison with the historical scenario.

Figure 7 - The caption says the graphs are showing mean spawner abundance, but the y-axis label says median spawner abundance. These values could be different, so ensure that the labels are correct and consistent with the caption.

Figure 8 and text on Page 29 - The procedure for repeating the exact sequence of historical ocean conditions is not clear for the warm scenario given that all plots start with ocean conditions in 1950. Were just the years 1988-2007 repeated sequentially after 1987 in the warm scenario? Also the caption does not specify whether the A1B or B1 emission scenario is shown (although the text says there is not much difference).

Figure 9 - It would be helpful to explain why some quasi-extinction probabilities are well above zero, even when the plots in Figure 7 do not show any mean abundances <50 spawners.

Figure 10 - it would be worth noting in the caption that the plot for “Big” is essentially uninformative because the cumulative probability was already very high under the historical scenario (evident in Figure 9) and, as stated in the text, must by definition go up in scenarios that consider later years.

Page 21 - Response metrics - delete “below” in the reference to “eq. 2, below”

Page 37 - It is reassuring to see that competition and predation are acknowledged as important parameters driving survivorship, but there are many other potential factors (e.g., interactions with non-native species, disease, contaminants). They should be acknowledged too.

2.5: Life-cycle matrix models to evaluate productivity and abundance under alternate scenarios for steelhead populations

1) Clarity of model goals

The objectives of the model are not clearly stated, although in the Introduction, the authors describe the focus of this chapter, "*In this document we will describe development of life-cycle models for O. mykiss applied to interior Columbia River basin populations,"* and describe its relationship to previous modeling by the ICTRT. It would be useful to add two or three sentences specifically identifying the goals of this modeling effort and how the model might be applied in the adaptive management cycle.

An additional question, not considered in this draft, is whether such models could be used to predict survival in the shorter term. Managers need quantified short-term predictions that can be evaluated with empirical data to demonstrate capabilities of the model and build confidence in model predictions. Model validation would be worthwhile in subsequent drafts of the model, if possible.

2) Soundness of methods and conceptual approach

The specific approaches and methods appear scientifically sound and are clearly written, but the models are still under development. Due to a lack of information about functional relationships, the scenarios comprise a set of fixed-level perturbations to parameters determining survival rates at various life stages. These perturbations may or may not include realistic values.

The modeling of life history variation seems somewhat deficient (preliminary) as resident life histories and iteroparity are not considered. Hatchery supplementation and the effects of hatchery/wild interactions are not included either.

The models also focused on populations at low elevation because, as noted by the authors, data are insufficient to model higher-elevation populations. Some populations were represented by "subunits" of the population because data are not available to model the entire population unit. Correctly delineating a population as demographically isolated from other such units is fundamental to a life-cycle model in which immigration and emigration are not modeled. The assumption that population subunits (that are only partially isolated from other subunits) will be demographically representative of the entire population seems highly questionable, and results could be misleading. The authors are commended for recognizing and acknowledging this subtle but important consideration.

The focus of the estuary/ocean component is restricted to early ocean survival for which only physical variables, i.e., water travel time in the mainstem Columbia River, fall PDO, and spring and fall upwelling indices, are considered. Modelers might consider incorporating biological indices and ecosystem indicators. Juvenile steelhead migrate offshore to the Gulf of Alaska during the first summer at sea. Thus, it might be expected that early (S_{O1}) ocean mortality of juvenile steelhead will have both a coastal and an offshore component that are influenced by different limiting factors. The ISAB encourages further exploration of other climate/ocean indices (beyond the PDO and coastal upwelling) that may be related to growth and survival of juvenile steelhead during the first-year offshore phase in the Gulf of Alaska (see comments on the Ocean Model).

Adult ocean survival is assumed to be constant (0.8 plus or minus 0.1 per year, citing Ricker 1976). Modelers might also consider later ocean effects, particularly, with respect to density-dependent growth and survival when Columbia River stocks are intermixed with abundant stocks of pink, sockeye, and chum salmon from other regions. These interactions during the first and subsequent ocean years may reduce growth and delay age of maturation, thus affecting ocean survival.

The authors point out that the "freshwater habitat" scenario is based on the assumption that improvements (or degradation) of freshwater habitat would simultaneously improve (or reduce) survival of ***both*** juveniles and adults by 10% during their freshwater life stages. This assumption implies a double impact, on two life stages, compared to the impact on single life stages in other scenarios, and so one might expect larger effects on spawner abundance compared to other scenarios. Certain kinds of habitat enhancement methods might benefit only juveniles or only adults in freshwater, but not both equally, because the two life stages might use different overwinter habitats (e.g., adults might need deep pools, whereas steelhead parr might need "river runs" with complex structure). Overall, the models appear quite sensitive to this parameter, which is unknown and difficult to estimate.

A key point is that, for some populations at risk, the increases in spawner abundance and decreases in quasi-extinction probabilities caused by improved freshwater survival of juveniles and adults (presumed to be possible via habitat enhancement) are sufficient to partly or completely offset the decreases in survival caused by recent climate change. This potential does not seem to exist for increases owing to improved downstream survival through dams or decreased avian predation, perhaps because of the assumed double benefit from habitat enhancement (as noted above). This potential double benefit effect should be evaluated.

3) Data use, availability, and gaps

These models would be more useful for management decisions if they included functional relationships between freshwater habitat conditions and juvenile and pre-spawning adult survival. What restoration actions might be needed to achieve the modeled survival rates? The authors note that efforts are underway to develop such models.

Life history diversity is not fully represented in the life-cycle model. Resident *O. mykiss* data are lacking for many populations, and the ISAB considers this to be a critical data need. Moreover, it seems that repeat spawners are not included in the model. Repeat spawners might be too scarce at present to contribute significantly to overall recruitment, but they do have a much higher fecundity than first-time spawners. Thus, it would be informative to use a life-cycle model to determine how much population viability is affected by kelt survival and repeat spawner abundance. Scenarios could be run to evaluate the benefit of enhancing kelt survival.

The Ricker (1976) ocean survival estimate used in the model was derived from data for sockeye, pink, chum, coho, and Chinook salmon. There is a critical need to determine the ocean survival of steelhead. Three scenarios are considered: historical (1946-2009), current or baseline (1985-2009), and "bad" PDO ocean (1977-1997). Some years overlap in the different scenarios, thus obscuring differences between years of good, bad, and neutral ocean conditions. In this sense the perturbations of ocean survival are less suitable for sensitivity analysis but do provide more realistic future scenarios than the arbitrary perturbations of other parameters.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

Readability is compromised by some awkward sentences and a tendency to repeat methods in the Results and Discussion section (see Specific Comments listed below).

The Results and Discussion section includes surprisingly little interpretation of the simulation results. The model is presented as being under development rather than ready for application; results are considered preliminary and subject to change.

As noted by the authors, the model results seem sensitive to starting values. The greatest uncertainties are values of S1 and outmigration/estuarine and early marine survival (s_{o1}). Is mortality of steelhead smolts migrating down the Snake River to the Lower Granite Dam ignored entirely in the model, or is it captured in the model somewhere other than in the s_{o1} phase? Time series of smolt and adult abundance data are short for some populations. Smolt estimates are uncertain for many populations. A major assumption is that the number of steelhead smolts is a repeatable function of the number of steelhead spawners, but this relationship might be greatly affected by resident adults (rainbow trout), which are not considered in the model.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

The level of complexity of the models seems appropriate given the data used. Development of more complex models will be required to resolve some of the conceptual deficiencies noted above.

The approach to sensitivity analysis, e.g., increasing or decreasing survival rates by arbitrary proportions (5%, 10%, etc.) is reasonable in the absence of known functional relationships linking management actions to changes in survival. However, the results presented seem preliminary because the sensitivity analyses are not comprehensive. A more systematic exploration of potential scenarios would be needed to evaluate the relative sensitivity of steelhead abundance and viability to changes in survival at various life history stages, and to assess which management actions might be most feasible for improving status. For example, which life stage should be targeted by restoration efforts in order to gain the greatest population benefit? Although many scenarios could be modeled, one of particular interest is the combined effect of all improvements in survival that could be effected by improved management, including for example, increases in survival owing to improved habitat, reduced harvest, improved downstream survival through dams, and improved estuary survival via reduced avian predation. It would be interesting to determine the extent to which such improvements due to management could compensate for decreased productivity due to changes in climate that could not be directly mitigated by management actions.

6) Adaptive management

Adaptive management was not addressed. It would be useful to add two or three sentences to the Introduction specifically identifying the goals of this modeling effort, and how the model might be applied in the adaptive management cycle.

Specific Comments

Page 5 (Modeled Scenarios) - The middle of the first paragraph is unclear; sentences are long and awkwardly constructed (clarity would be improved by ensuring clauses are parallel in structure). The scenarios described in the second paragraph do not actually represent changes in freshwater habitat but rather arbitrary changes in survival that are assumed to result (and might plausibly result) from unspecified changes to habitat. The same comment applies to the labeling of scenarios in first paragraph on page 9.

Why is stepwise multiple regression used to estimate survival in the estuarine-early ocean phase of the life cycle, rather than assessing a set of plausible models, or all possible subsets (e.g., see Burnham and Anderson 2002 on model selection and multi-model inference)?

Table 2.5.1 – It would be helpful if the caption referred the reader to text that explains how these survival values are calculated, and describes what time period they cover. For example, "B-H *a*" values are identical for 4 populations suggesting that the values are assumed rather than estimated; this should be clarified in the text on page 3. Over what time period is "estuary survival" calculated? The value shown for adult prespawning/overwinter "mortality" should presumably be relabeled as "survival" (unless 90% of adults die over winter).

Page 6 - The Results and Discussion section starts with: *"We updated the life-cycle models previously developed for the Umatilla River and Rapid River populations (ICTRT and Zabel 2007), extending the data series with seven recent years. We also successfully developed functioning models for six additional interior Columbia basin steelhead trout populations. For each population, we varied mortality rates at multiple stages of the life cycle to estimate resulting trends in population abundance and extinction probability over two time frames (25 and 100 years)."* These sentences would be better in the Methods section (or perhaps in the Abstract or the Introduction) rather than the Results.

Page 6 – The sentence at the bottom of the page that states "*these results show the potential range of the magnitude of changes in … abundance and viability*" is somewhat misleading. The results only show the magnitude of changes resulting from an arbitrary range of assumed perturbations.

Some of the tables would be more useful if they were depicted graphically in terms of "isoclines."

Chapter 3. Models Under Development

3.1: Snake River basin fall Chinook salmon run reconstruction as a basis for multistage stock-recruitment modeling with covariates

1) Clarity of model goals

The Introduction clearly describes why a two-stage model with covariates (hatchery production or other management actions) that can account for density dependence would be advantageous in assessing the effectiveness of management actions such as supplementation or habitat restoration. This chapter then describes the methods used to develop the abundance estimates for a two-stage model. The model will partition the life cycle essentially to above or below Lower Granite Dam (LGR). The above LGR stage requires reconstruction of the natural-origin adult run to estimate the adult return as well as the reconstruction of juvenile passage to estimate the number of smolts produced in the next generation.

2) Soundness of methods and conceptual approach

The text is methodical in taking into account the series of complicating factors that must be considered in reconstructing estimates of natural-origin adult and juvenile abundance of fall Chinook at Lower Granite Dam. The figures are helpful, in general, in explaining the conceptual approach for the reconstruction, but the text lacks sufficient detail to evaluate the implementation.

For example, the reconstruction of the adult run is an accounting exercise using data on fish passage at LGR and expanding it to account for less than complete sampling and other artifacts (Figure 5). These expanded counts and information from CWT and fin clips are then used to apportion the adult run into components of natural, hatchery, or unknown origin. While the accounting exercise seems to be carefully thought out, it is presented using a "written description" rather than a set of equations. This makes it difficult to assess if the model is valid as the written descriptions are “approximations” of what was done. The accounting exercise for the juvenile run reconstruction is also presented in general form using written descriptions rather than equations. Again, this makes it hard to audit the model without knowing the internal steps.

More explanation of the equation for estimating collection probability would be useful. For example, it was helpful to read that a decrease in turbine allocation will decrease collection probability because decreased turbine allocation means increased spill. But wouldn’t a decrease in river flow decrease spill, in which case shouldn’t the terms river flow and turbine allocation have opposite signs in the numerator of the equation?

The authors indicate that the reconstructions have been programmed into an Excel spreadsheet. Excel spreadsheets are difficult for performing audits on the model, so the conversion to computer code (see *Last Steps* section) is welcome.

Methods of estimating confidence intervals (p. 23) will use a bootstrap, but it is not clear how all sources of uncertainty will be accounted for in the bootstrapping. For example, how will the proportions of hatchery fish marked with CWTs (used in the expansions) be bootstrapped? How will the missing window counts be bootstrapped? It is also not clear why an estimate of precision is needed when the estimates are used in a future model where process noise may overwhelm sampling noise.

3) Data use, availability, and gaps

It appears that a wide variety of data are used to reconstruct various runs that cannot be measured directly. It is not clear however, which data are most crucial for each step and the relative “confidence” in the various types of data. For example, the actual count in the 50 minutes is likely "correct." The expansion factors to account for times not sampled depend on accurate recording of when counting is done and not done – how well is this recorded? It would be helpful to create a matrix showing which data are used in which part of the run-reconstruction and the "confidence" in each type of data. If some data have very low "confidence," this would imply that additional effort should be expended to improve collection of this data type.

Can PIT-tag information be used to help estimate total counts for days with partial or missing daily information as well? For example, is there a relationship between PIT-tag counts and actual window counts that can be used to adjust either count and improve the apportionment? Or are several different relationships needed because of differing fractions of different runs that are PIT tagged?

Some funding (p. 25) was requested for radio tagging work, but it is not clear if this will provide estimates for the most critical parts of the reconstruction model. Some sensitivity analysis of the final model is needed to see which information is most critical for obtaining information about the run reconstruction.

The model as described makes it difficult to incorporate data that only indirectly provides information about the component parts. For example, page 10 describes how missing data on days without window counts will be imputed based on the previous and following two days, but other information, for example from other dams, may also provide information about the missing data. Bayesian methods would be ideal to incorporate multiple sources of information about a particular component.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

The Introduction and most of the section on reconstructing the adult run are clearly written, but other parts could be improved, especially the section on reconstructing juvenile abundance. For example, the first sentence of that section (p. 15) begins awkwardly. In the same paragraph, the definitions of CH0 and CH1 are confusing – yet critical to understanding the rest of the analysis. It seems that some subyearlings (CH0) turn out to be CH1s by the end of the sentence.

Because details of the bookkeeping are complicated, the text is often difficult to follow in places. If the chapter is intended to document methods and to make them repeatable by others, then equations are needed to provide greater precision. If that requirement is to be fulfilled elsewhere (e.g., through documents available upon request), then the text could be greatly simplified to improve readability and conceptual clarity.

Readability and clarity is also compromised by some missing labeling or explanation in figures (variables and units on y-axes of figures 3 and 4; abbreviations in the flow charts in figures 6-8 that have not yet been defined, e.g., SbyC).

The authors do a commendable job of identifying assumptions and uncertainties, although as yet (recognizing that these methods are still under development), they do not discuss the implications of these uncertainties or the sensitivity of reconstructed estimates to the assumptions. As noted earlier, it would be helpful to have a matrix showing which data are used in which part of the model and the degree of confidence in each piece of data.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

It is not clear from the report if a more complex model is warranted in place of a pure statistical approach where, for example, counts at LGR are regressed against the final adult counts and this simple relationship is "good enough." Adding complexity does not always improve a model's predictions, especially if the intermediate steps rely on data that is not well known or if the intermediate steps have substantial "noise" in the relationships.

Model outputs or sensitivity analyses are not described – we hope that results will be included in the document scheduled for May 2014.

6) Adaptive management

The Introduction sets the stage for future use in adaptive management, but this report only looks at the reconstruction of abundance above LGR (adults and juveniles). Presumably, the modeling below LGR (juvenile to adult) is done elsewhere? Even with this accounting, there is no development of the relationship between adults returning and juveniles produced as a function of program enhancement (e.g., more habitat produced). So this model appears to develop only the inputs to the actual model of interest.

Specific Technical Comments

p.12 Last paragraph - Why not just use sampling weights directly (e.g. 1/sampling fraction) rather than relative sampling weights? Using sampling weights directly would give estimates of the actual run abundance. In obtaining relative proportions, odd sampling weights (which are just scaled versions of the true sampling weights) would just cancel out (as expected).

p.22 Section 3.6 and 3.7 seem to have the same section title but actually refer to different time periods.

p.25. The Next Steps section requires updating (some dates have already passed or are imminent).

3.2: Methow River Intensively Monitored Watershed: incorporating food webs into the life cycle

1) Clarity of model goals

The goal and objectives are not clearly stated early in the report. It is difficult to determine what the goals are because jargon is used. The clearest explanation of the goals of this modeling effort came near the end, and these could be moved to the Introduction.

2) Soundness of methods and conceptual approach

This report is difficult to evaluate because it is not well organized. Much text is spent at the beginning describing monitoring efforts, which could be appropriate if some introductory text was provided about how the section would be laid out and how the pieces fit together.

The design philosophy was not directly connected to the actual models being used, although the philosophy does seem to be extensive with regard to selecting models. The figures are not helpful due to limited connection to the text, sparse captioning, and abstract representations. For example, Figure 3 is especially difficult to interpret because it seems like a figure designed for the project team and no one else. Who is the audience the authors are intending to inform with this figure?

There is little description of the Ricker model used and how it fits in with the other modeling being done. The Ricker curves seem to be fit with relatively little data (and no formal model selection was employed). As a result, much of the density-dependence assumed may rest on one or two data points in a small data set as seen in the Appendices. There was no discussion of the uncertainty in the Ricker parameters from these sparse data. There also appears to be confusion between the “average” carrying capacity (as measured by the Ricker curve) and stochastic variation around this point because of short-term weather and other events.

The last section on the ATP model should have been provided earlier. The authors mention that they have completed the periphyton biomass and production model, but no details are provided. For example, the authors should provide the equations constituting the model, describe its inputs and outputs, and describe the time step and spatial scale of the model. How was the conclusion formulated that salmon have a net negative effect on periphyton production (via a combination of bringing marine-derived nutrients but also disturbing substrate with their redd-building activities)?

3) Data use, availability, and gaps

Due to the lack of detail in this report, it is difficult to determine exactly which data are being used, how sensitive the model is to certain types of data, and if the data are being used appropriately.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

We recognize that the model is still in development. Nevertheless, the report is difficult to follow and the level of detail is inadequate to be useful in evaluating the models. A reader would not be able to replicate the findings. There is neither a discussion of assumptions and uncertainties underlying the models nor of the reasonableness of the model outputs.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

It is not possible to evaluate the models at this time due to the lack of detail.

6) Adaptive management

The description of how the models are developed and applied does seem to indicate an adaptive management approach. However, without a clear description of the actual models developed, it is not easy to envision whether the adaptive management approach is feasible. The iterative loop of modeling, monitoring, and evaluating and improving models is a key part of Adaptive Resource Management, and could be simply labeled as such.

Specific comments

There are multiple typographical errors in this section of the report.

Page 5, top: The order of the three guiding principles in the first paragraph doesn't appear to fit the order of the three in italics before it.

Page 6, Proposition 2: The last sentence is not stated well. The statement implies that the authors are looking for a quick and fast model to meet policy needs. It appears that they are trying to say that models are a tool to help analyze environmental conditions in a condensed time frame to assist with management and policy decisions.

Page 6, Proposition 3: What do the authors mean by "A model should be developed and validated independently?"

Page 7: Jargon such as "data harvester" should be defined or removed.

Page 7, Proposition 4: The second sentence is worded awkwardly and could be interpreted in several different ways. It makes it sound like one will know that the adaptive assessment of data and model integrity are finished when the model is not validated by field data, which is probably not the meaning intended.

Page 8, last sentence: Are the authors referring to Figure 5? How is the concept illustrated in Figure 5?

Page 8. Figure 4 and Figure 9 seem to be identical.

Figure 6: According to the caption, the figure shows planned and existing fish monitoring, but there is no way to distinguish what was planned versus existing in the figure. The yellow circles are very hard to see on white paper, and they are not explained (but there are a lot of them on the figure).

Table 1 is hard to follow. "Ricker p at .2k*/R^2" and "Ricker Redds @.2k" are unconventional symbols and are difficult to follow. Why do some columns have "(data)" in them? Should there be any units on any of the columns?

Page 11: Is Mullan (1992) the same as Mullan et al. (1992)? Citations should be more complete. This occurs on several pages but is not noted again.

Page12, bullet 5: The last sentence of this bullet is difficult to understand. These had been fixed as of when? Why does it matter if they had been fixed? What does "fixed" mean?

Table 2: What is the *** notation in the bottom part of the table referring to? The notation is not evident in the table.

Page 13: Question 2 discusses "out of basin effects," but this is the first mention of these, so a smoother transition would be useful.

Page 13: Toward the bottom of the page there is a reference to Appendix B but that does not seem to be included. The appendix may not be necessary for the purposes of this report.

Page 14: The database project is complementary to what?

Page 14: Define ATP the first time it is used. The reader must wait for Figure 8 to understand that it means Aquatic Trophic Productivity.

Page 15: There is another reference to Appendix B here (is this the same one as before? This one seems like it might be more relevant).

Figures 10, 11, and 12: These figures need much more explanation. The arrows appear to obliterate some of the text. It is unclear what the figures are trying to convey. It is unclear what the boxes labeled $C_{freshwater}$ and $P_{freshwater}$ refer to.

Page 19, top: The authors could also refer to Naiman et al. (2012, *Proc. Nat'l Acad. Sci.*), which includes a summary of the information in ISAB (2011) and is published in a refereed journal.

The relevance of the items in Appendix A is unclear. It seems like this information might be quite relevant to the model development and verification but does not seem to be referred to in the main text. Reference is made to Appendix B, but none was provided in the report.

Appendix Tables 1 and 2 - is Ford et al. (2010) listed in the References?

Appendix Table 3 - It would be good to clarify that the entries are the number of juveniles estimated to have smolted at each age and that the entry farthest right in each row is the total smolts divided by estimated eggs. Likewise, it is unclear whether the bottom right entry is the average listed below it (instead of to the left, as one might expect). Would it be better to add an additional column for smolts/redd, and then give the average for both smolt/egg and smolt/redd at the bottom?

Appendix Figure 3 - Make clear in the caption that data for the Columbia ecosystem are from Quinn (2005).

P 29 - It would seem desirable to use formal model selection procedures to test whether the data warrant a unimodal curve vs. simply a line to support the statements made about density-dependence. Many of the quadratic curves seem to be highly dependent on a single point for the curvature. The quadratic relationship for Chinook may be suitable, although those for steelhead seem less likely, the text notwithstanding. Is Table 8 of Appendix A the same as Table 3? Why is it presented twice?

3.3: Catherine Creek life-cycle model with policy optimization

1) Clarity of model goals

Goals for the population model and the policy optimization components are clearly stated, but the goals for the habitat model are not discussed explicitly in this section.

The model is very ambitious – Catherine Creek will be divided into distinct habitat sub-units where spawning and rearing occur. Fish in the sub-units are allowed to move between sub-units during rearing (which will require a large number of mixing parameters to be estimated). When the adults return, they can spawn in the different sub-units, which will again require a large number of mixing parameters to be estimated.

Each unit will have a Beverton-Holt relationship with influences by temperature, flow, fine sediment, and total spawning areas (more parameters).

They will assume that productivity and capacity are predicted by habitat conditions. A stochastic model will then be used to predict how changes in habitat affect productivity. This will then be used to see how allocation of effort among different activities leads to the "best use of resources."

2) Soundness of methods and conceptual approach

The conceptual approach is clearly described at a high level and is conceptually sound. Likewise, the population model uses standard approaches and is clearly described. The policy optimization model is under development, so few details are provided. The authors correctly note that theoretical relationships need to be empirically evaluated to determine if the modeled connection between policy choices and environmental variables are valid.

However, insufficient detail is provided to judge how well some important and challenging issues will be dealt with. For example, how will "best outcome" be defined given that there are "a variety of spatial and temporal constraints of resource availability"? Also, recovery outcomes under alternative scenarios are likely to satisfy different objectives and values to various degrees. How will these multiple accounts of value be combined to assess the "best outcome"?

3) Data use, availability, and gaps

The authors claim to have some prototype models developed, but no details are presented. Models need to be developed to predict, for example, the relationship between tree canopy height, shade, water-velocity and stream-temperature. These need to be also empirically calibrated.

Data requirements are not discussed, but as noted in the Introduction, they are extensive. It appears the data are not yet sufficient to make effective use of the model, and, as the author's state, more data are necessary to evaluate the realism of the policy optimization model. It is not clear from the description if the necessary data are already collected and simply need to be accessed and used, or if new data collection efforts are needed. It seems that data available from ISEMP and IMWs could be effectively used for model calibration.

Once the available data are available, "Prototype theoretical relationships … need to be empirically calibrated for better realism, both functionally as well as spatially." How will this be done?

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

The report is clearly written but the discussion remains at a simplified, conceptual level. Methods are not described in sufficient detail to judge their merit, let alone to repeat analyses. More discussion is needed on how spatial and temporal variability will be incorporated into the existing deterministic models. At least a preliminary evaluation of the accuracy and precision of the policy optimization output should be possible with existing empirical data. The uncertainties and weaknesses in the analysis are acknowledged, but the consequences are not discussed.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

The level of complexity seems appropriate to the objectives, but no model outputs are presented. The first sentence of the second last paragraph states, “Early prototype simulations have produced results indicating that it is feasible to define … and accurately determine optimal allocation of restoration activities ...” How was this accuracy determined? At this point the modeling is “in progress” so sensitivity analyses have not been completed.

6) Adaptive management

The stated objective for the model is to explore alternative measures for population recovery. As such the model will serve as a tool for adaptive management. It would be beneficial for the authors to provide some illustrations explicitly describing how they see the model being used for adaptive management.

3.4: Yakima River steelhead and other *Oncorhynchus mykiss* populations

1) Clarity of model goals

The goals and objectives of the model are clearly stated. The primary focus is to understand the population dynamics of sympatric populations of resident (rainbow trout) and anadromous (steelhead) *O. mykiss*, in particular (1) abundance trends over time and (2) relative proportions of anadromous vs. resident individuals produced in the main stem and main tributaries of the Yakima. The model uses a wealth of information available on many life history parameters for these fish.

A secondary goal of the model may be to assess the effects of flow modifications from the various dams on the relative proportion of steelhead vs. rainbow trout in the future, but this is not clearly stated in the Introduction.

2) Soundness of methods and conceptual approach

The model is not described in great detail in this report because the fine details are covered in several cited papers, but the overview is clearly written. A main conceptual flaw is acknowledged, but nevertheless could be very important. The conceptual flaw is that the model lacks any function to describe the effects of environmental factors on the relative number of migrants vs. residents. As a result, it is not possible to suggest (as the authors do later) that the difference in life history tactics among upstream mainstem habitats and downstream tributaries is owing to these environmental factors, because they are confounded with the spatial variation in these factors. Any other factor that differs among the tributaries, which is correlated with the differences in temperature and flow, might also explain the pattern.

Also, in this model, the fish's decision to be anadromous vs. resident is genetically determined. The authors suggested that the model needs further development to incorporate environmental factors. In addition, changes in migration and ocean survival, harvest, and adult overwintering survival in freshwater are not considered and need to be incorporated into the model.

Could the relationship between smolting and temperature be hump-shaped? Here it is stated that warm temperatures that result in poor growth cause fish to smolt and produce more steelhead, compared to cooler temperatures that favor less smolting and more resident rainbow trout. However, if temperatures were too cold for growth of juveniles (e.g., perhaps with hypolimnetic releases, downstream from some reservoirs), would anadromy again be favored as growth declined? Perhaps temperatures would rarely, or never, be this cold in this river system.

3) Data use, availability, and gaps

The first model uses a 15-year dataset (1992-2006) of abundance, age structure, life history tactic, and maturation status. The second model uses information from nine spatial units in the Yakima River with information on habitat availability, stream flow and temperature, growth, dispersal, etc. These datasets are used to estimate many (but not all) of the parameters of the model – some parameter values are derived from the literature.

As noted earlier, data and relationships to sort out the effects of environmental factors on smolting, which apparently are available in another set of models (see Discussion), are needed to build a better model. This is not to say that the present model is not useful, but the authors realize that a key attribute may be missing.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

This section is clearly written and methods are described in sufficient detail, although it might be difficult to replicate results without access to model code and input parameter data files. A major assumption of the model is that smolt-to-adult survival is based on size at emigration and migration distance. The authors might also consider the effects of timing of ocean entrance, sea surface temperature at ocean entrance, and length of residence in the Columbia River estuary and plume as potential major factors affecting SARs.

The authors indicate in the Discussion that another report addresses the limitations of the model, but then appear to go on to summarize them here. Were all of the limitations summarized in this report?

5) Model complexity, usefulness for comparisons, and sensitivity analyses

The model is not described in great detail because the fine details are covered in several cited papers. For the most part, the complexity of the models seems appropriate.

However, it was unclear how food availability is incorporated in the second model as it is unlikely that data will be available at the appropriate level of detail. The model includes no influence of environment on the decision to smolt or mature in freshwater. Why? Is this because we have no good data on this relationship?

If the second model lacks any function to describe the effects of environmental factors on the relative number of smolts vs. residents, then how can the model suggest that the difference in life history tactics among tributaries is owing to these factors? Isn't this simply inferred from the spatial distribution among tributaries, correlated with the differences in temperature and flow? This is especially important because the next paragraph makes a sweeping statement about how changing flow regimes via irrigation will alter steelhead production. Would the evidence for this stand up under scrutiny, given the large economic impact of this decision? The authors recognize this limitation and argue for development of a more complex life-cycle model for interior Columbia River steelhead that incorporates functional relationships between environmental conditions and freshwater carrying capacity and between environmental conditions and fish survival and growth (as developed by Courter et al. 2009). Additional model components that incorporate functional relationships between early ocean conditions (particularly sea surface temperature) and fish survival and growth are also needed.

The authors did perform some sensitivity analysis on their models. For example, the SAR and proportion of steelhead produced by crosses appear to be confounded. Additional sensitivity analyses are needed to see what other relationships are confounded, e.g., construct a matrix showing the relationships of responses as one variable at a time is changed.

6) Adaptive management

The role of the model in adaptive management was not identified but understanding how flow management affects production of anadromous and resident forms is an implied priority.

Specific comments

The title needs to better reflect the content of this section. For example: Modeling the dynamics of sympatric populations of resident (rainbow trout) and anadromous (steelhead) *Oncorhynchus mykiss* in the Yakima River and other interior river populations.

Various errors in spelling and grammar throughout could be fixed with judicious editing. See for example the bottom of p 12. What is a “resident male female offspring”?

Page 2, l. 12: It is not clear what “reducing the thermal profile” means, nor how this favors resident fish.

Page 5, middle: Standard fisheries terms are used to denote age here (e.g., age 0), but these don’t match the coding in Chapter 1.

Page 10, first full paragraph: In the first sentence, is this rate for resident rainbow trout only, not anadromous?

Page 10, under Results and Discussion: Does the steelhead model in Chap 2.5 include adult overwinter survival? If so, could it be used here?

P 10 bottom: Is the coefficient of determination the only, or the best, measure of model fit? Might mean square prediction error be a better metric? The figures here are of low resolution/quality and could be improved.

Page11, Fig 3.4.4: The fitted relationship is shown, but it would seem useful also to show the data used to fit this relationship and a measure of the precision of the fit.

Fig 3.4.9: Could this plot be misleading, because the temperature changes on the X-axis are relative to the observed mean, not an absolute standard, and this observed mean varies among sites?

Page 17, top: A summary of model limitations should be presented here, rather than forcing the reader to find another report to learn about those, given their importance. Or, are these limitations what follow? If so, this should be clearly stated so the reader knows.

Page 18, second last paragraph, first sentence: To what does “it” refer?

Chapter 4. Hatchery Impacts

4.1: Impacts of supplementation on population dynamics of Snake River spring/summer Chinook salmon

1) Clarity of model goals

The manuscript provides an excellent introduction to the problem. The primary focus is to develop a cost-effective modeling approach to estimate population-dynamic parameters of hatchery and wild salmon in mixed populations, i.e., intrinsic productivity and capacity. This objective might be expanded to say that this novel approach is used to evaluate the influence of hatchery supplementation on wild salmon conservation.

2) Soundness of methods and conceptual approach

The specific approaches and methods appear to be scientifically sound and are clearly written. Choices about assumptions and consequent limitations are discussed candidly and insightfully. There were no significant conceptual flaws identified, but the ISAB does raise below a number of specific issues that should be considered by the investigators.

The model is sophisticated. Because only aggregate data are available, it is extremely important that contrast in the wild- and hatchery-spawning proportions be available at both high and low spawner densities. Table 1 shows contrast in the pHOS values, but it does not show if the contrast occurs at both high and low spawner densities. This may be the cause of non-identifiability noted on the bottom of page 5. We appreciate the discussion of the data contrast issue with regard to the anomalous positive productivity effect of hatchery fish shown in the best-supported LG model.

The hyperparameters (page 7) are modeled as being independent of each other for each population, but presumably, if a particular stream-cohort exhibits a lower than average wild-recruit per spawner, it may also exhibit a lower-than-average natural-recruit per spawner because of some common factor (i.e., the river is poorly suited to salmon). Given this, the following statement should be clarified: *"When wild and hatchery hyper-means are identical … the parameters in each population are identical as well."* For example, just because the MEANS of the hyper-parameters are equal (which forces the distributions of the hyper-parameters to have the same form), it does not follow that the individual parameters values sampled from the two distributions in the MCMC step are necessarily equal.

It is not clear why the common SARs are not also modeled using a hierarchical model, perhaps with some autocorrelation to account for multi-year events such a decadal oscillations.

There are several ways to compute DIC in multi-level models. While DIC is a measure of “fit vs. complexity,” this can take place at several levels. It appears that the DIC used in this study is the “lowest level” DIC which accentuates differences among models. Also, there was no discussion of goodness-of-fit (e.g., Bayesian p-values) to see if the model actually fits the data well, regardless of the DIC values. We wonder whether the DIC approach used here was a primary factor leading to the current top model, which had the unexpected finding of hatchery fish having higher intrinsic productivity than wild fish. Given the conflicting findings of the top model versus all other models with regard to intrinsic productivity, it is worthwhile to identify for the reader the approximate cut-off value for identifying less fit models when using change in DIC.

Perhaps the greatest weakness or uncertainty lies with the data used to parameterize the models (e.g., adult spawner abundance, pHOS, age composition, harvest rates, wetted channel widths, etc.). Additional information would be useful on the variation in methods used to generate these data for populations and time periods, if any, and resulting effects on model output. For example, during the earlier years of supplementation, fewer hatchery fish may have received marks needed to estimate the number of hatchery fish spawning in streams.

One assumption that warrants further consideration or explanation is that competitive ability among fish types (wild or hatchery) may be unequal but not asymmetric (page 5). It would be worth listing the types of behavioral interactions for which this assumption ***could*** be plausible. For example, the assumption seems implausible for competition between different weight classes of fish, where relative size (asymmetry) would determine the outcome more than absolute size.

The finding that Rmax is consistently lower for hatchery fish than wild fish is convincing and seems hard to dispute. However, the argument that this discrepancy stems from behavioral capabilities or genetic differences remains speculative. A simpler explanation would seem to be that hatchery-origin fish, lacking juvenile experience in the stream, return to different (probably lower) reaches and experience a locally higher density there than do wild-origin fish; in effect, hatchery fish may occupy smaller streams. This explanation could be tested.

The use of a common SAR value to back-calculate smolt production is an interesting approach, and it seems to be reasonably valid except for when in-river mortality, which contributes to the SAR estimate, varies with population, or when other key life history characteristics differ among the populations. Population-specific differences in the timing of smolt emigration and entry to the ocean and differences in pHOS among parents might also cause differences in SAR estimates that are not associated with common factors in the ocean. It is also possible that the populations have somewhat different distributions at sea. This common SAR approach shifts all of the variability in population-specific SAR values, including measurement error, to variability in smolt production. It seems that this approach will amplify actual smolt production, i.e., populations with below average SAR will have even lower smolt abundance, whereas populations with above average SAR will have even higher smolt production. Does this impact the findings when trying to look at each population? Cumulative effects of dam passage of SAR values might determine whether the current model could be applied more broadly (as stated on page 12).

Is the survival value Φ in the initial unlabeled equation specific to each population, or is the SAR averaged across populations as described earlier in the paper? Two pages later, it is clarified that it is the shared value common to all populations. This makes sense for survival at sea, but each population may have somewhat different in-river mortality if smolt counts originated from different dam areas, e.g., Tucannon versus Salmon River in this study. What was the common SAR for the natural-origin Chinook over time, and is this value influenced by availability of populations for a specific year?

Some Columbia River researchers, such as the CSS group, distinguish between SAR and SAS, where SAR only involves adults reaching the upper most dam and SAS is based on adults entering the mouth of the Columbia. The SAR values reported here were corrected for in-river harvests, so if appropriate, it would be worth clarifying that smolt-to-adult survival rates here are based on adults returning to the mouth of the Columbia River.

3) Data use, availability, and gaps

The investigation uses aggregate data on the total number of adults produced from the aggregate of the spawning population and requires high contrast in the natural- and hatchery-spawning densities. On page 12, the authors identify the need for contrast in the densities of both hatchery-origin and natural-origin spawners. Perhaps a graph showing the proportion of hatchery-origin parents over actual density would show if there are gaps in the data set. Some hint of this problem in the contrast is seen in Figure 1, where higher hatchery proportions tend to occur almost always at lower densities.

As noted above, some discussion is needed on the quality of the data. For example, total returning adults are adjusted to include in-river harvest (likely poorly known), aggregated across multiple ages (some estimates are better than others), and hatchery broodstock. Similarly, SAR values require estimates of juvenile outmigration which again may not be well known.

The last brood year used in the study is 2003, in which a four-year-old fish would have returned in 2007. Why were more recent data excluded from the analysis? Table 1 shows pHOS values for the 23 populations used in the study. Recent data provided to the ISAB from the HSRG indicated recent pHOS values were ~40% for Snake River spring Chinook, whereas the values shown here are typically much lower. If recent pHOS values are now much higher, please discuss how higher pHOS values may be affecting natural salmon. This topic is discussed indirectly when concluding the greatest possible benefit from supplementation occurs at low densities. But what does the model tell us about current management practices that have allowed high pHOS levels in some watersheds in recent years? Specifically, at what spawning densities should supplementation with hatchery fish be stopped, or slowed, in order to avoid replacement of wild fish with hatchery fish as stated in the discussion. Further linkage of the model findings to management practices would be worthwhile, though this application may be beyond the initial journal publication. Furthermore, evaluation of supplementation scenarios at high and low densities may vary with the model, given that the best-supported model shown here had a different finding compared with other models with regard to intrinsic productivity of hatchery fish.

What is the PNI value of the populations during supplementation, or what is pNOB?

The models make appropriate use of the data, but the authors note that they cannot rule out the possibility that their estimated differences between hatchery and wild fish are biased by unmeasured covariates unrelated to supplementation. In addition, observations at low values of S_h and S_w were reportedly insufficient to model effects of hatchery supplementation in the absence of density dependence. If so, what does this tell us about the higher intrinsic productivity of hatchery salmon identified in the best-supported model? As suggested by the authors, future analyses or experimental manipulations using data from depleted populations with a high proportion of hatchery spawners could provide further insights.

Additional experiments and data are needed to identify mechanisms that give rise to the differences in intrinsic productivity and carrying capacity between hatchery-origin and wild-origin fish that were inferred by the study.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

The model and results are clearly presented except it is not completely clear how supplementation effects in the supplementation scenarios are computed. The authors provide a balanced discussion of assumptions, uncertainties, and potential weaknesses.

The Figure 5 caption is confusing and could be improved.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

The level of complexity is appropriate given the limitations in the data. The models are just complex enough to estimate the effects of interest, yet remain simple enough such that data are available for comparing many populations with contrasting histories of supplementation. As noted above, there may be some concerns about the quality of some data, but it may turn out that natural variation in the values overshadow the data quality issue.

As the authors point out, more complex models would be needed to infer the reasons (genetic, behavioral, environmental) for observed differences in productivity and carrying capacity, but the data are not yet available to support such models. Accordingly, an experimental design to systematically manipulate the level of hatchery supplementation would be the most powerful way to determine the reasons for the observed differences in productivity and carrying capacity between hatchery-origin and wild-origin fish.

Informative sensitivity analyses were performed, e.g., Figs. 4 and 5. It would be worthwhile to elaborate more on why some populations had higher positive supplementation effects than others, if possible. For example, are mean values in Fig. 5 related to spawning density of each population, e.g., populations with low density have higher supplementation effect?

It would be interesting to see a three dimensional plot of change in hatchery and wild smolt production in relation to pHOS and total spawning densities, if possible. At what pHOS values and spawning densities do beneficial versus no benefit, or adverse effects, occur, based on the best supported model? Fig. 4 shows essentially no change in smolt production with increasing spawner density when the number of wild spawners was held constant at current median levels.

It seems that the model could be used to provide guidance on how many hatchery fish should be allowed to spawn in each watershed depending on numbers of wild-origin fish. An appendix with values or graphs for each population using actual spawner numbers rather than spawner densities would be relevant to managers.

6) Adaptive management

There is no explicit mention of adaptive management, but the supplementation scenarios inform what can be done to adjust supplementation programs to improve overall productivity. The authors also suggest that systematic manipulation of hatchery supplementation levels in an experimental design would be a powerful approach for learning about relative reproductive contributions of hatchery and wild fish across a range of population size. The Discussion includes insightful suggestions about the implications of the findings in this study. It identifies some new potential risks to supplementation for restoring populations and provides additional justification for habitat restoration given the significant effects of density dependence. These insights seem highly relevant to adaptive management within the Fish and Wildlife Program.

Additional comments

P. 10, para 2. Text says that Fig 2 shows productivity declining as proportion of hatchery. Should this be Fig. 1?

Text needs to be clear with regard to intrinsic productivity versus capacity and hatchery versus wild. Fig. 2 shows that hatchery fish have a higher intrinsic productivity than wild fish, on average, whereas the opposite is true for maximum smolt production – based on the top model. Text should say that intrinsic productivity was higher for hatchery fish whereas smolt production was higher for wild fish, and then should discuss the variability in estimates. The current text is somewhat confusing. The last paragraph on P. 10 is a much clearer explanation.

Fig. 1 suggests a steeper decline in smolts per spawner when spawner density increases and the proportion that are hatchery fish increases. Could this steeper slope for hatchery fish cause the higher intrinsic productivity of hatchery versus wild fish in the top model? It seems odd that no correlation was found between Ua and Up, but maybe this is because productivity did not vary much from population to population? Assuming the top model correctly identifies that hatchery fish have higher intrinsic productivity, is there any additional information to support this conclusion, or is it an artifact of the data or model type? Current thinking suggests hatchery fish would have a lower rather than higher intrinsic productivity, e.g., due to younger age/lower fecundity, domestication, possibility of hatchery spawners selecting lower quality spawning habitat, etc.

Discussion, 1st paragraph: The text states that the models favor lower values of both parameters for hatchery versus wild fish. But Fig. 2 clearly shows that intrinsic productivity is higher for hatchery fish and not lower (based on the top model). Clarification is needed because the top model, which is used in the figures, supports a different conclusion than all of the other models. What is it about this top model that would possibly lead to a different direction with regard to intrinsic productivity of hatchery fish? As noted above, we are curious if this top model is an artifact of the type of DIC calculation.

Fig. 4a identifies the overall spawning density at which hatchery supplementation may lead to an increase at low spawning densities versus a decrease at high densities. This should be discussed more. Can the model be used to identify a specific density at which supplementation switches from a possible benefit to no beneficial demographic effect?

What does the analysis tell us about spawning escapement goals for each population?

p.4. The last equation on the page should read E[R] rather than R.

Chapter 5. Estuary/Ocean

5.1: Estuary

1) Clarity of model goals

The goals and objectives of this section are clearly stated, i.e., to explore effects of a 50% reduction in estuarine bird predation on estuarine survival of Chinook and steelhead smolts. The effect on viability parameters is demonstrated in Chapter 2.5.

A related issue that was not addressed is whether predation mortality varies with salmon abundance or the abundance of alternative prey species, including marine species.

2) Soundness of methods and conceptual approach

The ISAB considers the text of this section too brief to determine the scientific soundness of the specific approaches and methods. For example, there is no explanation of how estimates of avian predation and estuarine survival are calculated. The authors simply state the values of parameters used and cite external sources for additional information. Nevertheless, the conceptual approach of parsing survival into estuarine and early ocean terms is important and necessary to many potential application of the life-cycle matrix model, e.g., to evaluate potential effects of estuarine habitat restoration. For example, information in Chapter 2.1 (p. 42) indicates that future Grande Ronde Chinook population models will include separate "estuary and early ocean components if warranted based on progress in ongoing analyses." This will have to be done if researchers plan on examining the effects of habitat restoration in the estuary.

3) Data use, availability, and gaps

As noted by the authors, their use of the data is only illustrative. The authors recognize that more data are needed, in particular, information on year-to-year variation in estuarine survival of steelhead and Chinook salmon smolts and other (non-avian) sources of estuarine mortality. The results in Section 2.5 (P. 182, Life-cycle matrix models for steelhead populations) indicate that "increases in estuary survival resulting from decreased avian predation and increases in survival through the Snake River hydropower system resulted in smaller percentage changes in the median number of spawners across populations than did the climate and habitat change scenarios." While bird predation is clearly a major source of mortality, the life-cycle model does not take into account a number of other possible factors such as estuarine habitat changes, contaminants, and food web effects.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

This section of the report is unusually brief. The level of detail in the methods is likely insufficient for replication. The authors only briefly discuss assumptions and uncertainties (that is, no compensatory mortality; constant estuarine survival and predation rates).

5) Model complexity, usefulness for comparisons, and sensitivity analyses

The life-cycle model was modified to include an explicit estimate of estuarine survival of Chinook salmon and steelhead smolts. Whether this level of complexity is appropriate for the life-history models cannot be evaluated given the limited information in the report. At present the avian predation and survival parameters used are not population-specific. Sensitivity analyses are not conducted in this chapter or in Chapter 2.

6) Adaptive management

The role of adaptive management is identified and integral to the goal of achieving the 50% reduction in bird predation. However, details of how managers might accomplish this goal are not considered.

Editorial comments

The title of section 5.1 needs to be more informative of the content of this section, for example: Effects of a reduction in avian predation on Chinook salmon and steelhead smolts in the Columbia River Estuary on population parameters and performance.

Some of the figure and table numbers in section 5.1 are incorrect.

5.2: Ocean conditions

1) Clarity of model goals

The goal of this chapter section is to provide details on modeling of ocean survival for spring/summer Chinook and steelhead populations. The ICTRT and Zabel (2007) analyses for Snake River spring/summer Chinook salmon, Wenatchee spring Chinook salmon, Snake River steelhead, Umatilla River steelhead are updated and new relationships for Yakima River steelhead are developed.

2) Soundness of methods and conceptual approach

There is no scientific justification for the methods used other than to cite previous work by the ICTRT and Zabel (2007). The variables chosen seem reasonable, although there is little information as to why they were chosen and no hypotheses are presented supporting their use. There are potentially significant conceptual flaws in the model related to the complex ocean life histories and ocean distribution and migration patterns of Columbia River steelhead and Chinook salmon. For example, juvenile Columbia River steelhead migrate rapidly northwestward to the open ocean (Gulf of Alaska) shortly after ocean entry. Thus, 1st-year ocean survival (S_{O1}) of steelhead includes both local (coastal) and distant-water (Gulf of Alaska) components, and appropriate indicators of survival during these two phases likely differ. Columbia River Chinook salmon exhibit diverse, population-specific distribution and migration patterns during their 1st and subsequent years at sea. The few variables evaluated might or might not be the best indicators of ocean conditions in the regions where Chinook salmon migrate. Similarly, it is reasonable that there are distant-water and local (coastal, estuary) components to survival of adult salmon and steelhead in the year of return to the river. While broad-scale indicators like the PDO may be related to both local and distant-water ocean conditions, they may also be strongly related to freshwater and estuary conditions. Thus, in the current model the PDO might not be an indicator of 1st-year ocean survival.

3) Data use, availability, and gaps

The available data seem adequate, and there is a large array of studies supporting underlying hypotheses for the use of these and some other variables (although not presented in the paper). Nevertheless, the indicators used generally have a stronger relationship to Chinook salmon survival than to steelhead survival. In particular, the steelhead models might benefit from further evaluation of indicators. For the local (coastal) S_{O1} phase, the authors might consider relationships between body size at ocean entrance, timing of ocean entrance, sea surface temperature at ocean entrance, and length of residence in the Columbia River estuary and plume as potential major indicators of ocean survival. The authors might also evaluate local biological indicators such as the copepod, forage fish, predator, and salmon growth indices being used or developed to forecast marine salmon survival in the Northern California Current system (e.g., www.nwfsc.noaa.gov/research/divisions/fe/estuarine/oeip/index.cfm).

For the distant-water component of S_{O1} of steelhead, further exploration of climate/ocean and indices (beyond the PDO, coastal upwelling, and Northern California Current ecosystem indicators) is recommended. For example sea surface temperatures and prey availability in the Gulf of Alaska are potential indicators of survival.

Ocean survival is assumed to be constant (0.8 ± 0.1, Ricker 1976) after the first ocean year. The Ricker (1976) ocean survival estimate is derived primarily from sockeye, pink, and chum salmon data collected during the pre-1977 regime shift period. There is a need for species-, stock-, and ocean life-stage specific estimates of ocean survival such as those determined by acoustic telemetry to validate this assumption.

While the authors seem to consider the Ricker (1976) estimate sufficient for their purposes, model performance might be improved by incorporating species-, population-, and life stage-specific estimates. Indicators for later ocean life-history stages could be developed and evaluated. For example, time series of growth indices from scales have proven to be useful indicators of survival in other applications. In addition, ecological interactions with other species and stocks (including hatchery/wild interactions) may reduce growth and delay age of maturation, thus influencing ocean survival. Abundance indices of other salmon species and stocks known to intermix with Columbia River salmon and steelhead in the ocean might also be useful indicators of ocean survival.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

The results are very preliminary, and the strengths and weaknesses are not fully described. The AIC approach to choosing among models seems to be a fairly mechanistic approach that warrants more ecological interpretation. That interpretation is not provided in this report.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

The use of a best three variable model selection is probably appropriate and may be adequate for many purposes. Nevertheless, adding additional complexity to the estimates of ocean survival and consideration of additional indicators of ocean conditions may improve explanatory and predictive capabilities.

6) Adaptive management

Adaptive management is not considered in the section of the report.

Chapter 6. Hydrosystem Survival

1) Clarity of model goals

The authors indicate that the existing model for estimating hydrosystem survival will be used to model future scenarios for hydrosystem operations and improvements, as well as scenarios of water withdrawal under the Columbia River Treaty Process. No details are presented on how the CSS and COMPASS models will be used to model dam breaching and reservoir drawdowns.

2) Soundness of methods and conceptual approach

It is not possible to determine if there are conceptual flaws since the context of the modeling is not clearly presented in this chapter. The material presented is too brief to ascertain how effective the modeling effort will be. Assessment of the approaches and methods depends on evaluation of CSS and COMPASS models. The ISAB did not attempt to review these documents for this evaluation. The ISAB review of the 2013 draft CSS Annual Report will occur soon.

3) Data use, availability, and gaps

It is difficult to assess the adequacy of the data because of lack of detail in the document. Although there are probably adequate data available to construct and model hydrosystem survival, the details are not provided in this chapter. The authors state that the data used to illustrate application of the model are placeholders to serve as demonstration of model capability until more detailed tributary habitat components become available.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

The application of the model is not described in sufficient detail to understand the capabilities and limitations of the model. Assumptions and uncertainties are not described.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

It would appear that one purpose of the model is to allow comparisons among populations across scenarios, but not enough detail is provided to evaluate whether the level of complexity is appropriate. There is no evidence that an assessment of the sensitivity of the model has taken place.

6) Adaptive management

Based on the material presented in this chapter it is not clear how the model would be used in adaptive management, although it would seem that adaptive management of hydrosystem operations is a probable application.

Chapter 7. Quantifying Spatial Structure of Interior Columbia Basin Salmon Populations

7.1: Introduction: toward a metapopulation model

1) Clarity of model goals

This entire chapter, of which this section is an introduction, is described as a work in progress. The implied objectives follow from the overall need to develop metapopulation models that identify populations at risk through isolation and help evaluate co-variation among populations articulated in the Adaptive Management Implementation Plan.

2) Soundness of methods and conceptual approach

This section is an overview, so no methods are described.

3) Data use, availability, and gaps

It is not entirely clear whether the report makes use of all the data/information available for the system considered here. The authors reference the unpublished work from the ICTRT, but not the published work in the Middle Fork Salmon that specifically addresses metapopulation questions and could be relevant here (e.g., Isaak et al. 2003, 2007; Neville et al. 2006, 2007). A more thorough review of the pilot system seems warranted.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

Not applicable.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

Not applicable.

6) Adaptive management

The potential role of the models in management decisions that might prioritize habitat restoration is discussed, but such an application is still only a possibility.

7.2: From genes to landscapes using multiple data sources to identify spatial conservation priorities for Chinook salmon in the interior Columbia River basin

1) Clarity of model goals

The Introduction is short, but the objectives of the paper are clearly stated as a series of questions to be answered through the analysis. More explanation of why these objectives are important would be helpful if the section is intended to stand alone, but the preceding section (7.1) provides some context.

The questions are important ones, to allow understanding the limitations of sources of data, and the sensitivity of analyses, which will be important in considering uses of the model for setting conservation priorities.

Given the uncertainty revealed, would it be better to state this (i.e., uncertainty) explicitly in the title as a goal of the analysis?

2) Soundness of methods and conceptual approach

The methods appear sound and represent new applications for the questions of metapopulation structure in Snake River salmon. The chapter clearly explains the conceptual basis for each of the models of dispersal and spatial structure. The alternative models have been developed methodically and logically by making a number of different simplifying assumptions. As the authors acknowledge, many of these assumptions are questionable, and the whole study remains a work in progress. The analytical effort is commendable, but in the end, the results are hard to interpret and unconvincing because they seem so sensitive to the assumptions.

Reviewers expressed concern about several points:

a. The choice of $p_c = 0.99$ for probability of a correct turn in the "wrong turn" model seems to have been arbitrary, and the results do suggest that this value is too low. If there is no way to estimate or validate this coefficient, is it really useful? The results suggest not.

b. Similarly, the choice of number of forks (T_h) seems arbitrary and subject to the fractal problem. The authors later acknowledge that size of tributaries should probably have been taken into account. It should be noted that the wrong turn model focuses exclusively on finding access to spawning habitat rather than on choosing spawning sites that are suitable for reproduction (because of local adaptation). This distinction is about homing behavior, and is additional to the concern that the wrong turn model does not take the requirement for actual reproduction into account.

c. There seems to be some confusion about genetic distance as the term is applied without much distinction to two different distance measures (CSE and Fst). The CSE distance is an appropriate measure of genetic distance for reconstructing phylogenies when gene frequencies have changed primarily through random genetic drift rather than mutation. In contrast, Fst (or more appropriately Fst/[1-Fst], as in the caption to Figure 6) is the appropriate measure to infer gene flow (mNe), because of the theoretical equilibrium relationship between these quantities. Note that mNe is the absolute number of immigrants into a population with effective population size Ne; m is therefore the (proportional) recipient-based stray rate.

d. The procedure for combining hydrologic distance and genetic distance seems rather *ad hoc.* Is weighting by genetic distance really the best way to combine these two independent estimates of isolation? Also what genetic distance index is used (CSE vs. Fst)? If each pair-wise genetic distance is normalized by dividing by the maximum pair-wise distance, then all but one of the normalized values will be less than 1, and the hydrologic distances will be reduced by the weighting. What then was done to refit the dispersal probability values to the weighted hydrologic distances (as done originally in Figure 2)?

e. It will be important to investigate the implications of having excluded populations (often many) for which data are unavailable. Presumably simulations with different number of populations could be used to test the sensitivity of results to the proportion of populations that are excluded.

3) Data use, availability, and gaps

Reviewers recorded several ideas and concerns about the use of available data to develop these models:

a. For the wrong-turn model, it is not clear why the value of 0.99 was assumed. If this value is simply a guess, or based on data, then that should be made clear.

b. The mark recapture data used to empirically estimate dispersal among populations is limited and subject to a number of difficult assumptions. These are clearly disclosed in the manuscript, but the end results are highly uncertain.

Overall, the authors have been careful to note where missing data limit or compromise their analyses. Ultimately, the different approaches provide very different pictures of the system, making it difficult to draw conclusions. The authors conclude that increased monitoring efforts are needed to improve understanding of dispersal mechanisms, but it is not clear that the critical assumptions of the models can be addressed to allow selecting among them for use in conservation.

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

We appreciated the care with which the assumptions and uncertainties were presented and kept at the forefront of all presentation of results and in the Discussion. The concepts are nicely illustrated with figures (e.g., Fig. 5) and substantial care is taken to explain the step-by-step process associated with each of the dispersal metrics. A section on the limitations and key assumptions is included in the methods for each approach. These provide a useful overview of the limitations, assumptions, and logic for each.

The intermediate results, parameter values, and management scenarios are extensively documented in tables. Not surprisingly, in so much detail, there are a few errors:

- On page 26, the authors need to better explain the source of (or method for deriving) the putative historical spatial structure for scenario H.
- Also on page 26, the authors should explain why hydrologic distance and genetic distance are combined as described.
- In Table 7, "variance" in column heading should probably be "change."
- On page 36 the references to CSE and Fig 6A are inconsistent; perhaps "Fig 6A" should be "Fig 6B"? But see previous comments about the use of CSE vs. Fst.
- In Table 13, "Table 8" should probably be "Table 9."
- Figure 8 is hard to comprehend and interpret. Is it worthwhile?

5) Model complexity, usefulness for comparisons, and sensitivity analyses

The authors had mixed success in modeling dispersal with the four approaches, and stated that a more complex model that includes other relevant covariates will be needed to accurately model dispersal.

Overall, the four different methods of dispersal gave rather different results when used to model strength of interactions, and without more data it is impossible to determine which method yields results that are closer to the truth. The two methods based on empirical data from the fish themselves, recaptures of coded-wire tags and genetic data, also give rather different results.

Given that the analysis presented here is likely in its infancy, the authors proceeded with appropriate caution in using these dispersal metrics to assess alternatives for conservation. However, it is unclear to us what values from which metrics are used for the "nominal values" for analyzing these alternatives (see Table 7) and how these relate to the different metrics for estimating dispersal/connectivity among the populations.

Overall, the most robust result is that the largest effects on dispersal and isolation are from increasing or reducing the influence of hatchery fish (Conservation Scenarios R1 and C1), but this is not highlighted in the Discussion (second to last paragraph). Another result that could be highlighted is the finding that subpopulations near mainstems are more closely related to neighboring populations than they are to subpopulations farther upstream in the same tributaries. This is the same finding as by Cooney et al., which was discussed in Chapter 7.1 above.

Overall, the different models are all useful from a conceptual standpoint, but results are not reinforcing. Each model explores a series of scenarios that might inform conservation management in the future. A crude sensitivity analysis is performed but without a clear understanding of the appropriate bounds for parameters. This is really a pilot analysis, useful for understanding the approaches and limitations of existing data. Considerable new work will still be needed to develop informative models that can be used by managers to compare or prioritize populations and actions.

6) Adaptive management

The authors attempted to evaluate a set of alternative future scenarios that could identify possible management and conservation options and their impacts on source-sink dynamics. Analyzing alternatives like this is a main component of Adaptive Resource Management. Models like these could clearly help focus management on critical populations, but that outcome will require better data than currently exist.

Other specific comments:

p. 2, end of penultimate paragraph - "… but which may be self-sustaining in the absence of …" Should this be NOT self-sustaining? That is, sink populations are not self-sustaining without immigration.

p. 6, bottom - Is there an objective statistical method for estimating the parameters in the dispersal kernel model, rather than estimating them by eye? Is there some reason that a statistical method would not be desirable? Estimating them by eye seems *ad hoc*, given the sophisticated methods of model fitting that are available.

p. 31 - "Clearly, the wrong turn dispersal model does not represent the way that fisheries biologists believe these populations interact with one another." Is this statement based on data, or another metric? This was unclear.

7.3: Spatial covariance of interior Columbia River spring/summer Chinook salmon from abundance data

1) *Clarity of model goals*

The goals are clearly stated: to use newly assembled time series of data to describe synchrony in the abundances of wild Chinook spawners across populations and to use this analysis to better understand the spatial structure of populations of this species. However, the link back to the larger life-cycle modeling framework is not clear. Is this work intended to inform the larger effort, and how might it do that?

2) *Soundness of methods and conceptual approach*

A primary concern is the quality and limitations of the data sets used for the analysis. The data are based on a variety of expansions or extrapolations that are vulnerable to a number of assumptions and confounding effects that differ across methods and populations. For example, many are based on index sites that may or may not represent local populations well and may change in that representation over time. Some are based on redd counts that have biases and unknown errors. Do these issues matter to the results?

A second concern is the development of the covariates based on dominance of 4-year-old fish. Although 4-year olds may be most common, younger and older fish are important and the age structure of adults may vary substantially within and among populations. This simplification may be appropriate given the coarse patterns of interest, but it would be helpful to make that argument. If age structure varies, considerable work and information would be required to decompose populations to the point that environmental effects could be linked to individual life stages. That would mean that the lags must vary by individual populations and that multiple year classes are associated with the annual adult estimates. It would be desirable to avoid this complex process if possible, so a more detailed discussion or some sort of uncertainty analysis is warranted to determine whether variable age structure is an important issue.

Specific comments:

a. Unfortunately, not all variables in equations 1a-1c are defined, and several are not defined clearly. For example, the variable *x* is never defined, and *y* is defined simply as "observations." As a result, the reader is never clear whether the standardized abundances of salmon are represented by *x, y,* or some other variable.

b. Many of the variance-covariance matrices used in these equations are defined using simplifying assumptions. Several are represented by the identity matrix (**Q**) or a matrix with equal values throughout (**R**). Is this justified, other than by the assertion that the model didn't converge without these simplifying assumptions? More explanation appears needed here.

c. Why is a forward stepwise procedure used for the environmental covariates, rather than developing a set of models *a priori* or all subsets? Could the order of entry affect which variables were found to be important, as can happen with stepwise methods?

3) Data use, availability, and gaps

Yes, the number of data sets with >50 years of data has increased recently, and this unique set of data has been put to good use with this sophisticated modeling. Importantly, the MARSS framework can account for both temporal and spatial correlations in the data, which are important to account for. The analysis seems appropriate and the data sufficient given the caveat about age structure (see comments above).

4) Clarity of descriptions of methods, assumptions, uncertainty, accuracy, and precision

Overall, the report is quite clearly written, which is especially important given the complicated modeling. In particular, the Discussion is well crafted and succinct – key findings are presented in topic sentences. The only exception is the issues with covariates and age structure described above.

5) Model complexity, usefulness for comparisons, and sensitivity analyses

There is some question of data quality and utility that might benefit from some sort of uncertainty analysis. That is, do the results differ if the potential errors in the extrapolated population estimates are considered?

6) Adaptive management

The role of the model in adaptive management is not identified clearly. The models do not directly support an analysis of population dynamics, and it is difficult to see how they might be used to refine more complex metapopulation models. Although the patterns are very interesting and potentially useful for understanding and conserving ESU level diversity, they do not seem to directly inform the metapopulation objectives that would support comparison of population dynamics or management actions. That is, the structure at the scales considered here may be less about metapopulation processes than about similarities in life history, or river or ocean migration patterns. The analysis does not really inform questions of synchrony and stability at the scale that populations are likely to interact in a demographic way. Some discussion of the ultimate application in the modeling framework anticipated by the larger report would be useful.

References

Arthaud, D.L., C.M. Greene, K. Guibault, and J.V. Jr. Morrow. 2010. "Contrasting Life-Cycle Impacts of Stream Flow on Two Chinook Salmon Populations." Hydrobiologia 655 (2010): 171-88.

Bjornn, T.C. 1971. "Trout and Salmon Movements in Two Idaho Streams as Related to Temperature, Food, Stream Flow, Cover and Population Density." Transactions of the American Fisheries Society 100, no. 3 (1971): 423-38

Burnham, K.P. and D.R. Anderson. 2002. *Model selection and multimodel inference: a practical information-theoretic approach, second edition,* Springer Science and Business Media, New York.

Courter, I., C. Justice, and S. Cramer. 2009. Flow and temperature effects on life history diversity of Oncorhynchus mykiss in the Yakima River basin. Cramer Fish Sciences, Gresham, OR.

Davis, J.M., C. V. Baxter, E. J. Rosi-Marshall, J. L. Pierce, B. T. Crosby. 2013. Anticipating stream ecosystem responses to climate change: toward predictions that incorporate effects via land-water linkages. Ecosystems 16: 909-922.

ICTRT (Interior Columbia Technical Recovery Team) and Zabel, R. 2007. Assessing the impact of environmental conditions and hydropower on population productivity for interior Columbia River stream-type Chinook and steelhead populations. Available online: http://www.nwfsc.noaa.gov/trt/col_docs/matrix_model.pdf

Isaak, D.J., R.F. Thurow, B.E. Rieman, and J.B. Dunham. 2003. Temporal variation in synchrony among Chinook salmon (Oncorhynchus tshawytscha) redd counts from a wilderness area in central Idaho. Canadian Journal of Fisheries and Aquatic Sciences. 60: 840-848.

Isaak, D. J. and R. F. Thurow. 2006. Network-scale spatial and temporal variation in Chinook salmon red distributions: patterns inferred from spatially continuous replicate surveys. Canadian Journal of Fisheries and Aquatic Sciences 63:285-296.

Isaak D, Thurow R, Rieman B, Dunham J. 2007. Chinook salmon use of spawning patches: Relative roles of habitat quality, size, and connectivity. Ecological Applications 17: 352-364.

ISAB (Independent Scientific Advisory Board). 2001-1. Model synthesis report. Northwest Power and Conservation Council. Report no. ISAB 2001-1. (20 July 2011; www.nwcouncil.org/fw/isrp/isrp2001-1).

Hilborn, R., T.P Quinn, D.E. Schindler, and D.E. Rogers. 2003. Biocomplexity and fisheries sustainability. Proceedings of the National Academy of Sciences, USA 100: 6564-6568.

Lucas, R.W., R.A. Sponseller, H. Laudon. 2013. Controls over base cation concentrations in stream and review waters: a long-term analysis on the role of deposition and climate. Ecosystems 16:707-721.

Naiman, R. J., J. R. Alldredge, D. A. Beauchamp, P. A. Bisson, J. Congleton, C. J. Henny, N. Huntly, R. Lamberson, C. Levings, E. N. Merrill, W. G. Pearcy, B. E. Rieman, G. T. Ruggerone, D. Scarnecchia, P. E. Smouse, C. C. Wood. 2012. Developing a broader scientific foundation for river restoration: Columbia River food webs. Proceedings of the National Academy of Sciences 109: 21201-21207.

Neville, H.M., D.J. Isaak, J.B. Dunham, R.F. Thurow, and B.E. Rieman. 2006. Fine-scale natal homing and localized movement as shaped by sex and spawning habitat in Chinook salmon: insights from spatial autocorrelation analysis of individual genotypes. Molecular Ecology 15:4589-4602.

Neville, H., D. Isaak, R. Thurow, J. Dunham, and B. Rieman. 2007. Microsatellite variation reveals weak genetic structure and retention of genetic variability in threatened Chinook salmon (Oncorhynchus tshawytscha) within a Snake river watershed. Conservation Genetics 8:133-147.

Roni, P., Pess, G., Beechie, T., and Moley, S. 2010. Estimating changes in coho salmon and steelhead abundance from watershed restoration: how much restoration Is needed to measurably increase smolt production? North American Journal of Fisheries Management 30:1469–1484.