Independent Scientific Advisory Board

Review of NOAA Fisheries'
Viable Salmonid Population (VSP) Modeling of Willamette River Spring Chinook Populations

(June 2014 draft)

ISAB 2014-4

August 1, 2014

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

ISAB Members

J. Richard Alldredge, Ph.D., Emeritus Professor of Statistics at Washington State University

Kurt Fausch, Ph.D., Professor of Fisheries and Aquatic Sciences, Department of Fish, Wildlife, and Conservation Biology at Colorado State University

Alec Maule, Ph.D., Fisheries Consultant and former head of the Ecology and Environmental Physiology Section, United States Geological Survey, Columbia River Research Laboratory

Kate Myers, Ph.D., Research Scientist, Aquatic and Fishery Sciences, University of Washington (Retired)

Robert J. Naiman, Ph.D., Professor of Aquatic and Fishery Sciences at University of Washington

Greg Ruggerone, Ph.D., Fisheries Scientist for Natural Resources Consultants

Laurel Saito, Ph.D., P.E., Director of the Graduate Program of Hydrologic Sciences at the University of Nevada Reno

Dennis Scarnecchia, Ph.D., Professor of Fish and Wildlife Resources at University of Idaho

Steve Schroder, Ph.D., Fisheries Consultant and former Fisheries Research Scientist at the Washington Department of Fish and Wildlife

Carl Schwarz, Ph.D., Professor of Statistics and Actuarial Science at Simon Fraser University, Canada

Ad Hoc Members

Robert Bilby, Ph.D., Ecologist at Weyerhaeuser Company

Chris C. Wood, Ph.D., Scientist Emeritus at the Pacific Biological Station, Department of Fisheries and Oceans, Nanaimo, British Columbia, Canada

Staff

Erik Merrill, J.D., Independent Science Program Manager, Northwest Power and Conservation Council

Cover photo: Detroit Dam spillway by Tony Grover, 2014 ISAB Tour

ISAB Review of NOAA Fisheries' *Viable Salmonid Population (VSP) Modeling of Willamette River Spring Chinook Populations*

ISAB Review of NOAA Fisheries' *Viable Salmonid Population (VSP) Modeling of Willamette River Spring Chinook Populations*

Background

This is the second of a two-part ISAB review associated with the Willamette Project, the U.S. Army Corps of Engineers' operated system of 10 high-head[1] federal dams and reservoirs, 3 run-of-river dams that function as re-regulating projects, and 42 revetments located in Willamette River tributaries. The Bonneville Power Administration (BPA) is responsible for marketing and transmitting power generated from 8 projects, with the remaining projects being non-power producing facilities. The U.S. Bureau of Reclamation (BOR) administers a water marketing program for water stored in Corps' reservoirs to agricultural users.

The ISAB's two-part review covers the Fish Benefits Workbook and life-cycle modeling developed by the Corps, NOAA Fisheries, and other agencies to inform the Corps' Configuration and Operations Plan, which is to be completed by December 2014. This second review is of the life-cycle model document titled *Viable Salmonid Population (VSP) Modeling of Willamette River Spring Chinook Populations*, which assesses the biological benefits of proposed actions within the basin to listed Chinook salmon. The first ISAB review was completed June 23, 2014 and evaluated the Fish Benefits Workbook (FBW), which is designed to help evaluate alternative approaches to improving downstream passage at dams and associated reservoirs in the Willamette River Basin (ISAB 2014-3).

The ISAB's reviews are intended to provide constructive feedback to the Corps, NOAA Fisheries, and their cooperators as they complete analyses supporting the Configuration and Operations Plan. This review is a logical next step from the ISAB's recent review of NOAA's life-cycle model for the Federal Columbia River Power System Biological Opinion (FCRPS BiOp) (ISAB 2013-5) and the ISRP's Review of the Research, Monitoring, and Evaluation Plan and Proposals for the Willamette Valley Project (ISRP 2011-26). The Configuration and Operations Plan's work relates to reintroduction of anadromous fish to areas blocked by high-head federal dams – an issue that is being discussed basinwide, especially regarding the Columbia River Treaty.

The Corps, NOAA Fisheries, and agency representatives participating in the Willamette Action Team for Ecosystem Restoration (WATER) helped develop six questions for the ISAB's review. The questions are based on NOAA Fisheries' questions for the ISAB's 2013 review of the life-cycle model for the mainstem Columbia and Snake basins. For this review, the ISAB restructured the questions to accommodate the format of the life-cycle model document. The following section includes the ISAB's responses to three questions that address substantial issues about the overall document and is intended for a general audience. The final section

[1] "High-head" – The elevation change from forebay to the average tailwater ranges from 360' to 450' feet and reservoir water surface elevations fluctuate by as much as 160' annually at WP dams, creating significant challenges for developing fish passage improvements.

includes more specific comments on individual chapters and is intended to help the authors and investigators improve the report and refine future modeling efforts.

ISAB Comments on the Overall Life-cycle Model Document

The ISAB commends the authors for their sophisticated effort to grapple with a very difficult task. Developing a life-cycle model of this complexity for Willamette Chinook populations is particularly challenging because the population-specific data are insufficient to parameterize the model with confidence. In short, the modeling effort is far ahead of the empirical data needed to fit and verify the models. Consequently, model outputs (e.g., spawner abundances and VSP scores) are likely unreliable, and the absolute values should be interpreted cautiously. The ISAB also suspects that uncertainty has been underestimated such that the confidence intervals on outcomes are probably too narrow. The relative values of model outputs might still be useful for ranking alternative dam management options within rivers, but it seems questionable whether the final VSP scores actually differ significantly among scenarios given the wide confidence intervals on the scores. These concerns are discussed in more detail with respect to the following three questions that address conceptual rigor, empirical support, and reliability of predictions, respectively.

Q1. The Configuration and Operations Plan (COP) to address ESA concerns in the Willamette River will be based on Fish Benefits Workbook (FBW) and Species Life-cycle Analysis Modules (SLAM) modeling outputs to estimate VSP scores under alternative management actions. Are the SLAM and VSP modeling approaches scientifically sound? Are there any significant conceptual flaws?

The approach for estimating VSP scores based on outputs from the FBW and SLAM models appears to be conceptually and technically valid but cannot yet be deemed scientifically sound due to the lack of data available as inputs for the models. The process used to generate scores is very qualitative for two of the VSP parameters (diversity and spatial structure), but as explained in the report, this qualitative approach is necessitated by a lack of data to conduct a more quantitative assessment. Although the procedure used to measure life history diversity seems plausible and appropriate for present purposes, it is not yet possible to verify whether juvenile life history diversity is an adequate index for tracking overall life history diversity. Still, some consideration of diversity is better than none. Similarly, the modeling of domestication effects through a reduction in egg-to-fry survival as a function of pNI (proportionate natural influence) is somewhat speculative, but the approach seems plausible and consistent with current theory, and we commend the effort to consider genetic impacts of hatcheries.

Further clarification is needed about how stochastic variability is represented in the model. In particular, it is not clear which components of the model include year-to-year variability and which do not. Judging from the details provided in Chapters 2, 4, and 5, it appears that a constant transition fraction with no stochastic variability is used to characterize many of the transitions between life stages (e.g., from eggs-to-fry). It seems that only a few sources of variability are included in the model, most notably the Beverton-Holt relationship between the number of spawners and the number of eggs laid, and some of the transitions within the Cougar reservoir. Year-to-year variability in the size and age of returning spawners is not included. Consequently, it appears that the output trajectories from the model will underestimate the actual variability in the system. This underestimation implies that the confidence intervals reported for the VSPs measures will be too narrow.

The modeling of age at emigration does not seem to take into account the density of young salmon or how density might affect emigration patterns. For example, the proportions of juveniles emigrating as fry, sub-yearlings, or yearlings might be strongly influenced by density in rearing habitats. The study of gill ATP-ase (Romer et al. 2013) suggests that most juveniles passing dams are actively migrating seaward, but this pattern might not hold under higher rearing densities and slower growth rates. Density dependent patterns of juvenile migration might explain the high variability observed in life history traits such as age at maturation and age at smolting as inferred from adult scale or otolith samples. If this conjecture is true, then the index of diversity based on juvenile life history proportions may not adequately represent or track overall genetic diversity that constrains life history diversity in these populations.

Q2. Are the models adequately supported by empirical data? If not, what types of data would need to be collected? Do the models provide an appropriate framework to incorporate new data over time throughout the implementation of the BiOp?

The authors have worked hard to incorporate data and results from scientific studies into the modeling effort, and we commend them for these efforts. Individual parameter values were first based on literature reviews of studies done in the Willamette basin or in a similar system. Because of deficiencies in available information, parameter values were then adjusted based on expert opinion or workshops to match the observed data. Finally, the SLAM model was tuned again to match the baseline conditions. In the end, it is unclear from the documents just how much the final parameter values have been adjusted from empirical values in the literature or how much tuning was needed to match existing baseline data. Many combinations of parameters might lead to the same matches, so more discussion is needed about which parameters were adjusted and why.

Model predictions with the existing data should be regarded as highly suspect. The most serious deficiency in the modeling process is the lack of sufficient, location-specific data. Gaps in empirical data should be specifically identified and prioritized in a table or appendix. The primary gaps appear to be related to the distribution and survival of juvenile fish in the North Santiam River.

Faced with an absence of empirical data, the authors have made a commendable effort to parameterize models with ranges derived from expert opinion generated through formal procedures at workshops, rather than arbitrary guesswork. However, this process would be more credible if the participating experts were identified and the variation in their opinions documented in an appendix. In particular, it would be informative to report which parameters created the greatest divergence in expert opinion; these highly uncertain parameters could then be subjected to more intensive sensitivity analysis.

The modeling and sensitivity analysis in this report should help identify what kinds of additional data need to be collected to enable more quantitative assessment of the effect of dam management options on spatial structure and diversity. Indeed, it is very evident from the modeling so far that pre-spawning mortality is likely to have a major impact on the potential for developing self-sustaining populations of Chinook salmon above the dams in the Willamette basin. Much more field research and model evaluation is needed to adequately explain and predict pre-spawning mortality. (See our detailed comments on pre-spawning mortality under Chapter 2 in the following section “ISAB Comments on Individual Chapters”).

Given that the purpose of this model is to evaluate the effect of various dam management options on population performance, another critical data gap is the lack of information on fish survival and residence time in the reservoirs. It will be important to improve the understanding of how reservoirs are used by juvenile Chinook and steelhead, the predation impacts they experience there (which likely vary among reservoirs), and the overall effect of dam operations on egg-to-smolt survival. The stocking of rainbow trout to support recreational fishing in the reservoirs raises additional questions about potential ecological and genetic (in the case of steelhead) impacts on the restoration of anadromous runs. The FBW does not address any of these issues directly because its focus is restricted to estimating passage survival from the forebay to the tail race of each dam. The SLAM model does include estimates of survival rate within reservoirs, but the information on which these estimates are based is clearly incomplete. A concerted effort is needed to collect data to better understand how the reservoirs affect growth, migration, and survival of salmon and steelhead.

Better data on habitat types and quality, both above and below the projects, would also be of great value. Information on stream length accessible to fish, availability of spawning habitat, water temperature, and dissolved gases is included in the model, but essentially no quantitative data are included about other aspects of habitat condition. Estimates of habitat capacity in the current model are based on the availability of spawning gravel, with production adjusted based on the presumed effects of temperature or dissolved gases. However, the availability of rearing habitat, cover, or food all might have a significant effect on survival and growth of the fish after emergence from the gravel, yet these factors are not explicitly considered in the SLAM model.

The model structure is flexible enough to enable the incorporation of many additional components for evaluating dam management options as new data becomes available. However, as has occurred on the Snake and Columbia rivers, it is very possible that habitat restoration efforts to improve habitat conditions above or below the projects may ultimately be considered as a method to mitigate for survival impacts associated with the dams. Most habitat restoration projects focus on improving habitat attributes that are not currently included in the model (e.g., large wood placement, off-channel habitat enhancement, nutrient supplementation). As a result, the current model would have to be modified to assess the contribution of these types of restoration efforts to fish population performance. The modeling team should consider the potential benefits of strengthening the manner in which the model addresses freshwater habitat quantity and quality.

Q3. Is the level of complexity of the models appropriate? Have appropriate sensitivity analyses been conducted? How reliable are the predictions of outcomes and relative ranking of status under alternative mitigation scenarios?

As with the FBW, there is a tension between model complexity and data needs. For many parts of the SLAM model, complete data are not available for the populations being modeled, so information has been generalized from other populations or estimated by expert opinion. Individual components of the model were tuned to match data that are available for the study population, and then, the overall model was tuned again to match recent baseline conditions. Explicit documentation of the assumptions being made and of the sources of (or rationale for choosing) parameter values used in the models is needed to recognize the uncertainty in predictions, to communicate the appropriate use of the results, and to refine the models as more information becomes available.

The alternative dam operations being compared enter the SLAM model only in the module that predicts dam passage survivorship and dam passage efficiency (DPE). To the extent that scenarios in Figure 4.4 are identical except for assumed changes in DPE, it is not surprising to find an almost linear relationship between the predicted number of NOR (natural origin spawners) and the assumed DPE. This relationship could perhaps have been predicted without simulation modeling, at least while the current populations are so far below capacity. However, the more that proposed conditions differ from the current baseline, the greater the need for complex modeling, and for monitoring (or experimenting) to verify model performance.

Several alternative dam operations (e.g., selective withdrawal structures and delayed fill) are expected to change temperatures upstream or downstream of the dams. Although it is clear that operational effects on temperature regime were explored in a workshop (Appendix 2.8), it is not clear whether these expected temperature regimes are taken into account (as operation-specific input data) in calculating adult pre-spawning mortality and juvenile survival, which would affect the ranking of VSP scores across alternative operations. These temperature effects should be modeled separately for each option given that the major reason for the constructing selective withdrawal structures is to change temperature regime.

The Monte Carlo approach used in the sensitivity analysis of selected parameters seems reasonable, as does the normalization approach that allows the sensitivity of the different parameters to be compared. It will be important to provide further details about how parameters were selected for inclusion in the subset of parameters evaluated in the sensitivity analysis and to discuss how sensitive outcomes are to initial values used in simulation runs. (See comments on Chapter 6 in the following section “ISAB Comments on Individual Chapters.”)

The existing model is probably too unreliable to be used to predict absolute outcomes given limitations on the quality of data currently available. It is especially disconcerting that so little is known about parameter values for the North Santiam River (see page 7.6). Consequently, estimates of VSP scores or abundance of NOR are not likely to be reliable, and we believe that the confidence intervals around the VSP scores are underestimated. However, the model might still be useful to compare and rank relative outcomes under different management scenarios within individual populations. The authors clearly recognize the limitations of the model as currently parameterized and conclude that "VSP scores should be viewed as relative metrics of the different passage/operations alternatives assessed compared to baseline conditions." However, given all the variability (which is probably underestimated) around VSP scores, it seems questionable that the final scores differ significantly among the various scenarios being compared. (See comments on Chapter 7 in the following section "ISAB Comments on Individual Chapters.") Given all the uncertainty, it might also be true that self-sustaining, natural origin runs cannot be achieved under any of the scenarios considered. In any case, the analyses should be repeated and revised as data are collected to fill gaps, allowing the modeling effort to be refined as part of the adaptive management loop.

ISAB Comments on Individual Chapters

CHAPTER 1

1.A. Is the chapter clearly written? Are the methods described in sufficient detail for a reader to understand and replicate what was done?

Chapter 1 provides a clear and concise overview of the purpose and strategy of the VSP modeling.

The second guideline under Productivity and Abundance criteria (page 1.5) seems misstated. How can a growing population have an average abundance equal to the historical average? If the population is growing, the average must also be growing, in which case a time frame for computing the average should be specified.

1.B. Are existing data and scientific knowledge used effectively and appropriately?

Not applicable.

1.C. Are assumptions and uncertainties about the analyses clearly described? For example, do the authors identify the strengths and weaknesses of the model and its inputs, and accuracy and precision of model output?

This chapter indicates that a key component of uncertainty in the model involves the survival and migration of juveniles in the reservoirs, for which relevant data are unavailable. It seems that this uncertainty is not adequately represented in the model or in the report. After the model was initialized with parameters from the literature, the model was extensively tuned to match the observed data. However, it is not clear which parameters were adjusted. There are likely several possible combinations of parameters that could be tuned that would lead to the same baseline results.

Chapter 1 (or chapter 7) could be improved by including a table that describes each parameter and its uncertainty and indicates the sensitivity of the model predictions to the parameter. Such a table would highlight key data gaps and help to prioritize future research as well as assist the reader in understanding the reliability of the VSP scores.

1.D. Specific Comments and Editorial Suggestions

In Figure 1.1, it would be helpful to indicate the major facilities or structures that are relevant to the models (e.g., dams, reservoirs, and monitoring sites referred to in later chapters). Alternatively, a figure for the appropriate region that shows these locations would be useful at the beginning of Chapters 4 and 5.

CHAPTER 2

2.A. Is the chapter clearly written? Are the methods described in sufficient detail for a reader to understand and replicate what was done?

Chapter 2 describes the data available to parameterize the model. For each section of the SLAM model, the background literature is briefly summarized and the main parameter values are discussed. Although Chapter 2 adequately covers the studies and literature that justify the approach and choice of parameter values used to develop the SLAM model, it is not always clear what assumptions or parameter values are actually adopted for the model. The organization of the chapter also makes it difficult to review and assess all of the assumptions and inputs. Appendices are included without clearly explaining which values or elements from the appendices are used in the model. Several of the appendices read as if they are meeting notes, with useful documentation of expert discussions, but no clear indication of what elements of the appendices are included in the models. Each section also indicates that some tuning was done to match the SLAM model output to observed values, but the description of the tuning is brief and it would be difficult to reproduce the steps taken. In short, the document would be more convincing if this chapter were re-organized to highlight the assumptions, their rationale, and how they are applied in the models. Bulleted summaries at the end of each section might help to describe the section's relevance to the SLAM, FBW or VSP modeling.

It might also be worthwhile to reverse the sequence of chapters 2 and 3. Chapter 3 provides a broader overview of how the model outputs are combined to produce VSP scores, and provides context that might help a reader to follow the specifics in Chapter 2. Also, the diagram in Appendix 2.7 (labeled Figure 1 on page 2.114) is very helpful in understanding how SLAM handles survival and juvenile migration, and it would be worth referring to earlier.

The writing could use some proofreading to remove repetition and sentence fragments. Brevity and clarity are compromised by describing the various background studies first, rather than just stating what is assumed, and then citing references to support the assumptions. Some parts of the document are confusing due to editorial errors - for example, there are two tables labeled 2.2 and two labeled 2.3, and the wrong Eq. Box numbers are listed on page 2.121 (1 and 2 should be 3 and 4, respectively). Also, the axes of many figures show proportions mislabeled as percent, and vice versa. We have listed many such specific comments and editorial suggestions in section 2.D below.

The critically important section on pre-spawning mortality (PSM) involves a rather brief (4 page) description of an empirical model to predict PSM. Additional discussion is warranted (see our suggestions in section 2.B below). It is not clear whether the PSM values considered in this analysis apply to fish that are captured and transported around dams and reservoirs or whether some PSM values apply to migration before the dams are reached. This information is probably contained within the original data sources, but those references were not readily available online.

The equation (on page 2.7) for predicting PSM warrants more consideration. First, it is not clear why the logit(PSM) transformation was dropped in favor of untransformed PSM given that both models are essentially linear in the range of 20-80% and that more of the variability in PSM can be explained after logit transformation (r2=63% versus 40%). Without transformation, it seems possible to generate negative values; do the PSM values predicted by the linear equation require further adjustment to constrain values between 0 and 1 (as shown in Figure 2.3)? It appears that the logit transformation of PSM was discarded because it indicated 100% mortality in some reaches in most years. Why is this result unreasonable? Were temperature measurements in the lower river not representative of what Chinook experienced? If so, were the same data used in the linear models? Alternatively, did the raw data indicate exceptionally high PSM in some watersheds, i.e., perhaps supporting the logit transformation? The authors note “Another issue in developing a predictive PSM model is that many of the time series for PSM included operational changes in water releases from the dams and changes in fish handling procedures below the dam.” Can these changes be tested and potentially incorporated into the model? Also, what is meant by the statement, “Equations were often too sensitive to temperature changes near the mid-point”? Does this mean that the best-supported models do not adequately capture the relationships found in the raw data?

Second, it would be useful to describe in more detail the statistical analysis used to test the alternative sets of variables. Both fish density and % wild are reportedly correlated with PSM, in addition to maximum river temperature. However, because density and % wild are themselves correlated, it is uncertain which variable is influencing PSM. Moreover, it is not clear how density was calculated, because it was not evaluated further. Apparently, the variable % wild was included because it seemed to reflect the number of hatchery fish aggregating near the hatchery collection sites “where PSM was especially high.” More investigation is needed to determine whether PSM is related to fish density and/or the percentage of wild fish in the sample group. At present, it seems that % wild is being used as a surrogate for density, yet the density metric was not examined. To address this issue, the authors could test whether PSM is higher for hatchery versus wild Chinook salmon when both are transported around dams. They could also use density (or numbers) of natural and hatchery fish as separate independent variables. Note also that other investigators found that gender and time of arrival had a significant effect on PSM of Chinook in the Willamette (Keefer et al. 2010).

In the discussion of temperature effects in the North Santiam (pages 2.25 to 2.30), it would help to explain the objective of temperature control. Presumably it is intended to improve conditions for migration and spawning but the rationale in this section is very limited. Not all readers will have a good knowledge of the history of temperature control in these reaches.

The report indicates that some changes in dam operations affect water temperatures in the reservoir and downstream of the dam. However, it is difficult to discern from the documentation precisely how these temperature effects are captured. The documentation indicates that temperature is modeled to affect PSM (p. 2.9) and egg-to-emergent fry survival (p. 2.22), but more explicit documentation is needed on how these effects are implemented in the SLAM model. Moreover, it is not clear why only these life stages were selected for modeling temperature effects. Appendix 2.8 does include discussion of how temperature is affected by different operations, but the text and equations in that appendix are not clearly linked to what actually gets used in the models. A summary table showing where operational changes impact the SLAM model would be helpful.

In particular, the modeling of temperature impacts on PSM in the SLAM model needs clarification. In the discussion of temperature effects on PSM (starting on page 2.6), temperature is defined as the average maximum daily stream temperature for seven consecutive days. Why is the 7-day average of maximum temperatures used rather than a total exposure (such as degree days above 0 experienced by the fish) – or is such an exposure variable not feasible because SLAM does not model the time spent in the river as returning adults? The first SLAM diagram (Figure 4.2) shows an impact of environmental variables on spawners, but the variable is labeled as PSS A, which in turn is defined by a data file (MacKenzie_Temp_Max7-DADM.csv, page 4.29) that is presumably revised to account for temperature changes due to operational changes. The second SLAM diagram (Figure 5.2) also shows an impact of an environmental variable on spawners, but the variable is labeled as Temp A, which is defined by a similarly named data file (Nsantiam_Temp_Max7-DADM.csv, page 5.33). Notation should be as consistent as possible among the components of the SLAM model. Both of these *.csv files indicate they represent 200 years of temperature data (based on a small number of USGS data years), but no explanation is given as to how the USGS data were expanded nor how they were changed relative to the baseline conditions.

The explanation of the effect of temperature on egg-to-emergent fry survival could also be improved. The effect of temperature changes caused by operational changes is “invisible” to the casual reader. For example, it took some effort to discern that under baseline conditions (from the Excel file) the proportion of eggs in Reach B that survive to emerge as fry is 0.452 (range 0.405 to 0.50), and that under the SWS option (page 2.120 in Appendix 2.8), this survival rate increases to 0.498 (range 0.471 to 0.525). (Note that there appears to be a typo on page 5.10 where the mean for the baseline scenario is given as 0.425, inconsistent with the Excel file). We presume that these higher values are entered in the Excel file for the SWS scenario, but we were unable to check as this file was not provided for review. Temperature effects on embryonic survival appear to be modeled in a similar way.

In general, the current format makes it extremely difficult for the reader to determine which parameters have been influenced by operational changes and how these changes are entered in the models. Are there operational impacts on important environmental variables other than on temperature, DPE, and TDG? TDG effects from operational changes also appear to be modified in a similar fashion with similar shortcomings in the documentation. These types of effects need to be clearly documented in the report, perhaps by providing a table to show how management operations change the baseline value of parameters and where these are reflected in component inputs.

2.B. Are existing data and scientific knowledge used effectively and appropriately?

Information from the background papers appears to be reviewed and extracted appropriately, and in general the existing information is used effectively. Many of the studies are old and based on conditions prior to dam construction, so they may not apply to modern conditions. In some cases, data used to parameterize the model are not entirely appropriate for the application, and "expert opinion" was frequently used to establish input values for the model. However, in these cases the authors include appropriate caveats.

Exceptionally high (e.g., 80-90%) pre-spawning mortality (PSM) has been documented for adult spring Chinook salmon transported into some tributaries of the Willamette Basin in recent years (Caudill et al. 2013). Thus, PSM could prevent the development of self-sustaining natural Chinook salmon runs in the Willamette Basin or could reduce population productivity such that few if any adults can be harvested. Investigators in the Willamette Basin recognize this critical problem, and research into factors contributing to PSM has begun (Schreck et al. 2013, Roumasset 2012, and other citations in Chapter 2). We would also like to see more discussion in the current document of factors that have already been examined in the Willamette, and we suggest some additional sources of information and ideas in the following paragraphs. We also encourage further investigation of factors affecting PSM so that actions might be taken to reduce this major source of mortality.

Pre-spawning mortality in the Willamette Basin differs among subbasins, and it varies widely among years within the same subbasin. For example, PSM has ranged from a low of 1% in the upper McKenzie River in 2008 to 95% in the Middle Fork of the Willamette River in 2007 (Roumasset and Caudill 2013). Two recent studies one by Schreck et al. (2013) and the other by Roumasset and Caudill (2013) evaluated the importance of environmental conditions and biological factors on the occurrence of PSM. Schreck et al. (2013) found that fish that died prior to spawning were heavily infested with parasites, possessed severe lesions, and were infected with multiple pathogens. It is hypothesized that elevated water temperatures experienced during the holding period are responsible for the heavy parasite burdens. Chapter 2 briefly notes the importance of pathogens to PSM, but other contributing factors are not discussed. The Roumasset and Caudill (2013) study found that warm water temperatures and fish density on spawning grounds were positively associated with PSM. They also discovered a negative relationship between the percentage of wild fish on a spawning ground and PSM. This negative relationship occurred because hatchery fish often assembled in large aggregations next to dams in low quality spawning habitat. Agonistic interactions among the fish at these locations may have increased secondary pathogen infestation rates due to abrasions or minor wounds that eventually led to high PSM rates. Other factors affecting survival may have been present. Few wild fish were found in these locations. Stress may be associated with the trap and transportation around the reservoirs, including high densities of the fish while in captivity.

Additionally, Schreck et al. (2013) highlight the complex interactions of migration conditions, particularly water temperature, energetic status, pathogen burden, and stress on PSM. They developed a model in which PSM is associated with migratory corridor duration. Caudill et al. (2013) showed that time spent in the mainstem Willamette is positively related to the distance between Willamette Falls and the tributary a fish homed to. They also found that migration speed (river kilometers traveled per day) increased later in the spawning season and is linked to increased water temperatures. Keefer et al. (2010, not cited in the document) reported higher mortality of female and early-arriving Chinook in the Willamette Basin. Other studies (not cited) outside the Basin (e.g., Bristol Bay sockeye: Quinn et al. 2007; Fraser sockeye salmon: Hinch et al. 2012; Puget Sound coho salmon: Feist et al. 2011; Yukon Chinook: Hamazaki et al. 2013) note the effects of spawning density, migration timing (early versus late arriving fish), and river temperature.

These findings and Schreck et al.'s (2013) model are consistent with speculation by the authors of the current life-cycle model, who state that PSM may be linked to the cumulative water temperatures fish experience as they migrate and reside in the Willamette River. We therefore suggest that the authors of the current Willamette life-cycle model examine Schreck et al.'s model to see if it might be a better predictive tool for PSM than the current two-factor model. If sufficient temperature loggers are in place, perhaps these, or similar data could be used to create a model that predicts the cumulative temperature units (degrees C above 0 over a 24-h period) that fish migrating to different subbasins in the Willamette experience at different times of the year. This cumulative temperature value could then be used as an independent variable in a linear regression to predict PSM, or it could replace the 7-day maximum temperature variable in the existing two-factor multiple regression model. The suggested procedure would give the modelers an opportunity to evaluate the effects of cumulative temperatures on PSM.

An alternative approach is to evaluate the extent to which PSM must be reduced to achieve self-sustaining Chinook populations above the dams given all the other factors affecting their survival. This alternative approach is appealing given the current limits to our ability to explain or control PSM.

The literature review of egg-to-fry survival rates is fairly comprehensive, but egg-to-fry survival rates vary so widely in the cited studies that it is difficult to justify or credit the 50-60% survival rate assumed in the models for river reaches with little or no habitat degradation. Egg survival might be strongly related to flow conditions during incubation, including flood events. River reaches containing degraded habitat will have correspondingly lower survival rates, but it is not clear how that lower survival will be determined. Fieldwork is needed to validate the egg-to-fry survival assumptions. It would be valuable to document flow regimes, measure gravel movement due to flood events, and determine percent fines in representative areas with and without degradation, as all of these factors are related to survival during incubation. The outcomes of such work could be used to refine egg-to-fry survival estimates in degraded and non-degraded habitat areas.

Additionally, the modelers point out that water temperatures experienced by fish prior to spawning may affect gamete quality and subsequent survival. They indicate that such effects will likely be difficult to detect in nature due to possible direct temperature effects on developing embryos. Thus, controlled laboratory experiments are probably needed to assess the effects of different water temperatures on gamete quality and subsequent progeny performance.

The spawning capacity estimates used in the model are substantially lower than some of the empirical observations cited in the literature. For example, how do the authors justify a 60% reduction in capacity (see page 2.17) from that estimated by Parkhurst et al. (1950)? Is this an arbitrary choice? The implications of this choice should be investigated by sensitivity analysis.

There is not much discussion of how temperature changes within the reservoir habitat would affect fisheries. Recent studies by William Connor and Kenneth Tiffan and co-authors on Chinook salmon in reservoirs and temperature impacts on survival in the Snake River Basin seem relevant; eight papers by these authors are listed in the References section of this review.

Equations are developed to predict mortalities in alevins and juveniles when total dissolved gas (TDG) exceeds 105% and 115% respectively, and mortality rates are discounted depending on month of the year. However, no data are presented to indicate the range of TDG that has been recorded in either basin. Moreover, the mortality rate calculation appears to be based on a single lab study in 1974. Other more recent studies by Mesa et al. (2000) and Backman et al. (2002) might suggest different values. The recent study by Backman et al. might be especially relevant as they examined gas bubble disease in juvenile salmon as they migrated.

2.C. Are assumptions and uncertainties about the analyses clearly described? For example, do the authors identify the strengths and weaknesses of the model and its inputs, and accuracy and precision of model output?

Most assumptions in the model are adequately explained, although they are so numerous that it is challenging to keep track of them all. The authors recognize the limitations of the existing models and provide appropriate caveats regarding the limitations of input data. Because of these limitations, they try to calibrate the SLAM model to match the observed output. However, the details of the calibration are generally vague, and often “ad hoc” adjustment factors are used to calibrate the model with the observed data. Calibration is mentioned in several places, often with reference to a baseline scenario, but it remains unclear how the various components and equations were calibrated, and to which data. It would be useful to dedicate one section to describing the calibration procedures in more detail.

In some cases, uncertainty is conveyed by showing a range of data values from the literature, but in other cases uncertainty is not mentioned. Greater detail in describing the data set would help communicate uncertainties in the analyses. It might be useful to include an appendix outlining data gaps and approaches to fill them, and providing a more detailed discussion of the uncertainties, precision, and accuracy of the models.

The r^2 for the linear equation used to predict PSM is only 0.40 indicating that more than half of the variation associated with PSM remains unexplained. For example, Figure 2.3 shows that when maximum 7-d temperature exceeds 20°C, PSM ranges from ~20% to 100%. At relatively low temperatures (<14°C), PSM ranges from ~0% to 70%. Is the variability in PSM (as seen in Figure 2.3) modeled in the SLAM model? (It is not apparent from the Baseline tab in the Excel spreadsheet.) It will be important for the life-cycle model to incorporate this unexplained variability when testing whether actions can lead to self-sustaining Chinook populations.

The authors have chosen to model spawner (egg) capacity as the primary variable creating density dependence in the life-cycle model. However, there is not much discussion to justify this decision. Why is density dependence associated with spawner capacity rather than other variables, such as limited habitat for rearing?

Hatchery impacts due to competition, predation, and disease transmission are not modeled directly but are said to be captured indirectly in the life-stage specific survivals. However, it is not clear how these effects have been estimated or assumed, and this is a real shortcoming in the overall consideration of hatchery impacts. Domestication effects are modeled based on plausible inferences from pNI such that the egg-to-fry survival of offspring of HOS is discounted by up to 60% whereas the survival of offspring of NOR with hatchery influence is discounted by up to 20%. Unfortunately, as the authors acknowledge, few of the populations are monitored adequately to provide reliable data on pHOS and pNI.

The current SLAM treats the basins independently of one another. Thus, it does not appear that the current SLAM model has the capability to deal with potential interactions among populations from different basins once they intermingle in the mainstem.

2.D. Specific Comments and Editorial Suggestions

Page 2.1 (and 2.53): There is exact duplication of text in several places indicating that the organization of certain sections could be improved. For example, on page 2.1, the text about juvenile life history monitoring is repeated word for word on page 2.53. The text on page 2.1 is specifically referring to Chinook salmon, but it is not clear that is the case on page 2.53, where text is about VSP scoring metrics.

Page 2.1: Figure 2.1 would be easier to read if the left axis were shown as a proportion. As it is, the bars for the less numerous juvenile Chinook >150 mm are barely visible.

Page 2.3: What is the basis for assuming that "the distribution of juvenile life histories emigrating from non-impounded areas was set, as a default, at equal proportions of outmigrant age classes"?

Page 2.4: Variability across calendar years in proportions of juveniles at each age might also reflect variability in adult spawning densities, and subsequently, juvenile cohort strengths. In other words, variability in proportions within calendar years may overestimate variability in proportions within brood years.

Page 2.5: Is there a reference or rationale for the statement “Until recently, ocean harvests would have had a strong effect on age structure.” What is different about the recent past? What does this mean for SLAM – is this statement implying that SLAM does not consider ocean harvest?

Page 2.7, text about PSM equation: It is unclear why a linear model is considered “more conservative.” How is “conservative” being defined (for example, does “more conservative” mean “more mortality”)? For the equation, how is “Temp” defined? (Also, on page 2.9, there is another reference to “temperature units” without a clear definition.)

Page 2.8. Perhaps it would be better to sort both tables by AIC_c scores. Also note the typo in the legend regarding the definition of Delta AIC_c.

Page 2.9: In Figure 2.3, pre-spawn mortality is plotted against temperature for three levels of the % wild covariate, but the data points appear identical for each plot, which does not make sense. Perhaps a better way to show the effects of the covariate is to use partial residual analysis.

Pages 2.9 to 2.19: does “capacity” in this section always refer to spawner capacity? When written as “spawner (egg) capacity,” is this implying that spawner capacity and egg capacity are the same (it does not appear so from the tables)?

Page 2.10: The reference to Table 2.2 is confusing – Table 2.2 on page 2.8 is about something different; a second Table 2.2 appears on page 2.13; subsequent table numbering will also require updating.

Page 2.10: More explanation is needed about why and how SLAM baseline scenarios were calibrated to contemporary run size by “identifying those downstream life stages where survivals have been dramatically reduced relative to historical conditions.”

Page 2.13 (Table 2.2): Comparison of estimates would be easier if the same units were used in both columns, rather than redds in one column and adults in the other.

Page 2.14: Is any cross-hatched area actually shown in Figure 2.3?

Page 2.14 (last sentence) and page 2.19 (last sentence about Table 2.4): typos (“we” should be “were”) and it is unclear if these are the spawner capacities that are used in SLAM.

Page 2.15 and 2.19: Tables 2.3 and 2.4 provide capacity estimates, but what is the form of the BH model?

Page 2.16. Different opinions are cited for the number of fish per redd (2 vs 2.5, etc.). What is the best information available, do these values change with density?

Page 2.17. It is challenging to follow the logic as to how these studies comport with Parkhurst's (1950) results.

Page 2.19: It seems that most of the fecundity data are specified as eggs per individual female, regardless of the size or age structure of the females returning. Is our impression correct? Are these the best available data for these fish?

Page 2.20: Rationale for using half the female fecundity to represent adult fecundity given equal sex ratio could be explained more clearly. It seems this assumption might have implications for survival (or consistency with other assumptions about survival) if, on average, males mature earlier than females.

Page 2.21: Table 2.5 is rather cryptic. What do the colors and the black box denote? Some description of the expert opinion workshops and reference to Appendix 2.6 would be appropriate here.

Page 2.23: The line drawn through data in Figure 2.6 must be derived from more than just the data shown. How is the lower limb derived?

Page 2.24: Reference to Figure 2.8 seems mistaken; presumably Figure 2.7 is intended.

Pages 2.24-2.25: If the San Joaquin River is warmer and the fish there are better adapted to warmer temperatures, is it "conservative" to use those data as the basis for the Willamette models? Figure 2.8 is difficult to understand.

Page 2.26: What is Figure 8?

Pages 2.25 to 2.30: The text about temperature control structures and selective withdrawal is confusing. How are "temperature control structures" defined? In tables 2.6 and 2.7, it seems that the first two rows refer to conditions before temperature control structures were in place, and the last three rows are after, but this is not immediately clear. The dates for the historical regime differ between Tables 2.6/2.7 and Figure 2.10, so it is not clear when temperature control was in place. (The footnote that refers to different dates in Tables 2.6, 2.7 and Figure 2.10 does not help to clarify the situation.)

Page 2.32: typo, missing "not" near end of second last paragraph?: "…they are [not] likely to be exposed to sustained high levels of TDG."

Page 2.33: It is stated that the impact of ocean conditions uniformly affects the different scenarios in SLAM, but this might not be true if ocean conditions and management scenarios both affect one or more of the three juvenile life histories differently. Ocean conditions experienced by juveniles at ocean entry (and the impact on early marine survival) will vary depending on migration rate and timing of ocean entrance (e.g., Scheuerell et al. 2009) and migration patterns (e.g., McMichael et al. 2013). These clearly differ by species, population, life history type, etc. How do the different management scenarios affect migration rate after dam passage and timing of ocean entry?

Page 2.33: What is the r^2 value for the correlation between ONI, PDO and average smolt to adult survival? The text claims only that these variables are “roughly correlated,” which is not obvious from Figure 2.11.

Page 2.35: typo last line “each” not “reach.”

Page 2.36: In sampling from the variability about the s3 relationship, were values drawn uniformly from within the 95% bounds denoted by the dashed lines in Figures 2.12 and 2.13 or from a normal distribution that corresponds to those lines? The former procedure would likely overestimate variability.

Page 2.40: Is the additional tuning over and above the individual tuning done for each stage of the life cycle as noted earlier in the chapter? Also, it would help to highlight the statement that survival estimates for life history stages in baseline-accessible reaches are parameterized to capture the effects of habitat conditions and ecological interactions with hatchery fish. This is an important point that could be explained more clearly.

Page 2.43-2.68: The section on “Parameterization of the VSP Criteria” could be split out as a separate chapter because this topic is about scoring metrics rather than describing how the parameters are set.

Page 2.45 (top): Because NOR and HOR gene pools are not isolated through generations, better to use the word deme instead of population in the phrase “…entering the natural [spawning deme’s] gene pool.”

Page 2.45: Why only a 60% reduction in fitness instead of the higher reduction (87.2%) suggested by Chilcote et al. (2011)?

Page 2.47: Why is it reasonable that any hatchery origin fish returning to spawn would be considered natural with no hatchery legacy?

Page 2.49: Presumably the VSP score based on abundance and productivity (i.e., P(QET)) would measure persistence in the face of normal environmental variation whereas the diversity and spatial structure scores would measure persistence in the face of systematic environmental change requiring adaptation, and unusual or catastrophic environmental variation.

Page 2.55: Table 2.13 lists juvenile life history stage as 1^{st} spring migrant, 1^{st} autumn migrant, and 2^{nd} spring migrant, which are different from terms used elsewhere (but probably easier to follow than other nomenclature—see comment below for pages 2.94 and 2.97). Also, the “>” sign before percentages seems incorrect (and/or is confusing).

Page 2.55: Could subsume “Fitness effects…” section (pages 2.43-2.48) within “Artificial Propagation” to avoid redundancy in the following pages.

Page 2.57: The broken lines denoting pNI isopleths in Figure 2.20 seem incorrect (except for the 0.5 value). For example, slopes of 2/3 and 3/4 should correspond to pNI values of 0.67 and 0.75, respectively, but they do not in this plot.

Page 2.61. Is spatial structure analyzed entirely outside of SLAM?

Page 2.63: What are units in Table 2.14?

Page 2.63-2.68: It is not clear how qualitative spatial considerations influence the VSP scoring or how variability is assessed for spatial structure scores. More explanation is needed. Presumably, the R/S ratio referred to in the caption for Table 2.16 is also the S/S ratio in the equation on page 2.67. If so, this notation should be consistent to prevent confusion.

Pages 2.92-2.112 (Appendix 2.6): The text here seems to imply that the authors are not modeling growth differently in the reservoirs. It is difficult to understand the slides on pages 2.101-2.112 – are the numbers in the tables the DPEs? Presumably the caveat “must sum to 1,” refers to addition within columns (not immediately clear).

Pages 2.94 and 2.97 (Appendix 2.6): The table-like inserts contain novel (potentially confusing) nomenclature (“Spring subyearlings passing the dam, entered the estuary (pre-smolt),” “Spring subyearling (Fry),” “Fall subyearling,” “Spring yearling”). What is the relationship between the various “fall” and “spring” fish? Are these the same as the fall-run and spring-run Chinook in the Columbia? How do these relate to the fry, spring subs and fall yearlings, or the 1st spring migrants, 1st autumn migrants and 2nd spring migrants? Furthermore, one may know that they passed a dam in the basin, but how is it known that they entered the estuary? Why would a pre-smolt enter the estuary?

Pages 2.113-2.116 (Appendix 2.7): Much of the text on page 2.113 is repeated two additional times on pages 2.115 and 2.116.

Pages 2.117-2.122 (Appendix 2.8): It appears from this text that temperature is accounted for in the models by changing survival. However, the rationale for changing the survivals is rather cryptic. The authors should provide a bit more explanation of the rationale – why is one option going to increase survival whereas another will decrease it? In the boxes for the equations, Box 1 has positive relative survival effect values, yet the survivals are decreased, whereas Boxes 2 and 3 have positive relative survival effect values, and calculated survivals are increased. For Box 2, the text says the relative improvement is 0.05 to 0.15, but the Box shows 0.05 to 0.25 and 0.05 and 0.25 are used in the calculated survivals. Where does the 0.25 come from? Why would operational temperature changes at Detroit Dam result in warmer summer adult holding temperatures but lower egg incubation temperatures for this scenario?

Page 2.121: The text indicates that “the magnitude of this change (decrease) in survivals is unknown.” What then is used in the model? Should we assume that none of the box equations apply to the Deep Drawdown with fish passage scenario?

Page 2.121: For the Glory Hole option at the bottom of the page, the text says survivals would most likely be similar to, or less than those in Equation Box 2. However, the Box 2 survivals are increased, whereas the Box 4 survivals are reduced – which seems to contradict the text.

CHAPTER 3

3.A. Is the chapter clearly written? Are the methods described in sufficient detail for a reader to understand and replicate what was done?

Chapter 3 provides a very readable description of how SLAM is structured and how the VSP scores are computed. Chapters 4 and 5 provide additional details that would allow readers to replicate the work.

Clarity would be improved by including more explanation of why and how the 500 model runs are performed – what is different about each run, and what is the same? Is this a Monte Carlo approach in which some parameters are being randomly selected from a range? If so, which parameters are randomly selected and which parameters are fixed?

3.B. Are existing data and scientific knowledge used effectively and appropriately?

Not applicable

3.C. Are assumptions and uncertainties about the analyses clearly described? For example, do the authors identify the strengths and weaknesses of the model and its inputs, and accuracy and precision of model output?

It is not clear if different QET (quasi-extinction threshold) values are used for McKenzie and North Santiam. Typically, a QET is used in models to denote a danger level at the limit of acceptable resolution in risk calculations, i.e., the level below which chance processes affecting extinction risk are not (or cannot be) modeled reliably enough to be useful. It seems odd, then, that the QET for North Santiam (150 spawners) is lower than that for McKenzie (250 spawners) given that the information used to compute extinction risk is less certain for the North Santiam than for the McKenzie basin.

More explanation and discussion is needed about how uncertainty in the VSP scores for diversity and spatial structure are computed.

3.D. Specific Comments and Editorial Suggestions

Page 3.3. A 5-year burn-in period was used, but other parts of the document say that a 20-year burn-in was used? This apparent inconsistency should be explained (or corrected).

Page 3.6: Abundance is measured as a geometric mean for the good reasons indicated. Falling below QET is defined as the population abundance dropping "below the QET threshold, on average per year, over a four-year period." This statement is not clear. Does it mean that the 4-year running average of abundance dropped below the QET at least once during the period of simulation (years 26-105)? If so, is the averaging arithmetic or geometric, and why? More explanation is needed.

Page 3.8, Figure 3.2. Clarity would be improved by re-labeling the x-axis as "P(QET)" or by indicating in the caption that "Extinction Risk" and "P(QET)" are being used synonymously.

Page 3.9. The equation for P(QET) is not a logistic regression as noted in the paragraph above.

Page 3.12. It is not clear how uncertainty in the VSP scores for diversity and spatial structure are computed. No details are given. From the text in Chapter 2, there does not actually seem to be an uncertainty for these quantities.

CHAPTERS 4 AND 5

4/5.A. Is the chapter clearly written? Are the methods described in sufficient detail for a reader to understand and replicate what was done?

Chapters 4 and 5 provide a detailed description of the SLAM modeling for the McKenzie and North Santiam basins. The flow diagrams at the start of each chapter effectively illustrate the complete model so that a reader can understand the modules and how they fit together. However, a reader would need to be familiar with SLAM (or have the SLAM manual nearby) to decode the conventions used in the SLAM diagrams. The accompanying Excel file gives the parameterization used for many of the transitions in the diagram. Even so, many data files and dynamic drivers are only briefly described in the document, and without these, it would be impossible to reproduce the models. Presumably, the total documentation will contain all of the driver files and the dynamic drivers (the "groovy" files).

The modeling of pre-spawning mortality (PSM) is not described in these chapters, which is curious given that the sensitivity analysis in Chapter 6 shows that PSM had the largest influence on the results. Although it is clear that operational effects on temperature regime were explored in a workshop (Appendix 2.8), it is not clear whether the expected temperature regimes are taken into account (as operation-specific input data) in calculating adult PSM (and/or juvenile survival). The Excel file contains data only for the baseline scenario, so it was not possible to determine what temperature data are used for other scenarios.

It would be helpful to highlight for the reader which components of the SLAM model include year-to-year stochastic variation and which do not. This important question can be resolved by studying the Excel files, but not without considerable effort.

Clarity would be improved by giving an example for each of the individual descriptions of the stage transitions. For example, how does the number of spawners determine the number of eggs deposited? A Beverton-Holt relationship is used to model density dependence, but details of the implementation are not easy to discern from the description. Similarly, when data files are used to model the transitions (e.g., for the ocean survival steps), the reader is left guessing what the data files actually contain because the data files are "hidden" from the descriptions.

It would help the reader to point out key differences between Chapters 4 and 5. For example, the number of reaches is different – 3 in the Chapter 4 and 5 in Chapter 5. Why do some reaches have variability in the egg-to-fry transition and some do not? (page 5.10)

An explanation is needed for why, in summarizing results for the 100-year trajectories, the geomean is calculated for NOR spawning abundance (Figures 4.4 and 5.4) whereas end of simulation values are used for pHOS and life history diversity (Figures 4.7, 4.9, 5.7, 5.9)?

4/5.B. Are existing data and scientific knowledge used effectively and appropriately?

This detailed modeling appears to capture all of the life history diversities known to exist in the populations. A comprehensive literature review and workshop were used to select appropriate parameter values. However, adjustments of these values in the tuning phase are not well explained.

Habitat capacity above the Blue River Dam in the McKenzie Basin appears not to have been considered. If not, why?

4/5.C. Are assumptions and uncertainties about the analyses clearly described? For example, do the authors identify the strengths and weaknesses of the model and its inputs, and accuracy and precision of model output?

These chapters (and Chapter 2) provide a detailed description of the modeling, but it remains challenging to keep track of all the assumptions made at each life history stage. A list of assumptions would be useful. For example, the survival rate from egg to fry is assumed to be constant over all years regardless of environmental conditions (see page 4.8) despite the discussion of the effects of water temperature on survival (Section 2.22). It is not clear why this effect of water temperature is not captured in the SLAM model.

The min/max values for capacity shown in the tables on pages 4.6 and 5.7 are far wider than would be seen assuming a normal distribution with the means and SDs given in the tables. This implies that the stochastic variation being modeled is less than the min/max in the tables.

Was a sensitivity analysis done on the initialization values on page 4.33?

4/5.D Specific Comments and Editorial Suggestions

Figures 4.2 and 5.2: Some guidance is needed on how to interpret these diagrams. For example, what is the "season" that is listed in each box? What do the different colors mean? (The SLAM manual indicates: green – stage that needs to be initialized; grey – stage that does not need to be initialized; light blue – environment variable; purple – dynamic driver variable.)

Pages 4.6 and 5.7: The calculation of egg production is hard to follow and should be clarified. It seems that the number of eggs deposited and available for hatch is not determined directly, but rather, that a Beverton-Holt curve is fitted for each reach in each year using a fixed productivity term (e.g., 2250 for NOR) and a variable capacity term (see Tables on pages 4.6 and 5.7); then, the BH curve is used to "look up" the number of eggs produced by the number of spawners returning in a particular year.

Page 4.34: It would help to clarify exactly what is varied in the 500 simulation runs for each alternative and to correct (or explain) inconsistencies with respect to the length of the runs. The text here says each run was 100 years long, but the text on page 3.3 says each run was 105 years long, and the INFO tab in the Excel spreadsheet says each run is 146 years long.

Table 4.3 (page 4.34) and Table 5.3 (page 5.40): It would be helpful to include the abbreviations to this table for the figures shown later in the section. What is the significance of 1571? Is that an elevation?

CHAPTER 6

6.A. Is the chapter clearly written? Are the methods described in sufficient detail for a reader to understand and replicate what was done?

The procedure for sensitivity analysis described in Chapter 6 is a standard application of existing methods and the explanation is acceptably brief.

Clarity would be improved by providing a more detailed rationale for the choice of parameters to be included in the sensitivity analysis. On page 6.4, it is stated that selection of parameters to include in the sensitivity analysis was a first step, and that two criteria are used to make this choice. In what way did the chosen parameters meet the two selection criteria? Because a complete list of parameters is not shown, a reader cannot assess whether other parameters should have been tested. There are almost 100 parameters in the SLAM model, but the sensitivity analysis tested only 10-15 parameters per model. It would be helpful to summarize (perhaps in a table) the reasons why certain parameters (e.g., egg-to-fry survival) were not chosen for analysis.

It is not clear whether the 500 simulation runs for sensitivity analysis described in Chapter 6 are the same as runs described in Chapters 3, 4, and 5? In Chapters 4 and 5 (pages 4.33 and 5.38) different initial values were used to start each simulation, but it is not clear whether each of the 500 runs in the sensitivity analysis started with exactly the same initial values. How sensitive are the model results to initial values?

6.B. Are existing data and scientific knowledge used effectively and appropriately?

The sensitivity analysis was conducted using a version of the model coded in R, which mimics the SLAM-version, except that ocean survival is modeled as a function of covariates.

The variables chosen should be expanded to include potential impacts of climate change. For example, how might changes in water temperature influence parameters other than PSM?

Good experimental design methods (e.g., highly fractionated-factorial designs) have been developed to investigate the main effects and two-factor interactions of multiple variables. These methods should be considered in lieu of the simulation methods because they can identify not just main effects but also important two-factor interactions between parameters, which the current sensitivity analysis fails to capture.

6.C. Are assumptions and uncertainties about the analyses clearly described? For example, do the authors identify the strengths and weaknesses of the model and its inputs, and accuracy and precision of model output?

The relative importance of parameters for each project alternative is estimated by calculating standardized regression coefficients (mean divided by standard error) for the predictor variables. This may not be the best measure of the sensitivity of a parameter as it confounds the magnitude of the response with the consistency of the response over other parameters in the analysis. To facilitate comparison, the standardized regression coefficients are further normalized by dividing by the largest standardized regression coefficient. This normalization step potentially assigns a low sensitivity to a parameter whose impact is related to the values of other parameters, i.e., where there is an interaction among parameters. Thus, while these methods are likely sufficient to identify the important main effects of parameters, they are unlikely to detect interactions.

6.D. Specific Comments and Editorial Suggestions

Page 6.4: The text says that three-way proportional splits for juvenile life history types were sampled randomly from a uniform distribution within specified ranges until a vector of three proportions summed to one. Were all three proportions resampled each time or just the last proportion resampled to satisfy the summing-to-one constraint?

Page 6.8 (and 6.13): Indicate the DPE for base and fss.lower runs shown in Figure 6.3 (and base, sws.weirbox.lowRE, delayrefill, fsc.750.weir, and fso runs shown in Figure 6.5).

CHAPTER 7

7.A. Is the chapter clearly written? Are the methods described in sufficient detail for a reader to understand and replicate what was done?

Chapter 7 presents overall conclusions from analyses of alternative modifications to dam operations on the North Santiam and McKenzie rivers. The section is clearly written but rather brief. As the concluding chapter, it should be able to stand alone for those that did not read the details in chapters 4 and 5. The chapter could be improved by including a short introduction to provide continuity with the rest of the document – one that briefly restates the objectives of the modeling effort and the various management options being evaluated. It could be strengthened by presenting (again) the final VSP scores (e.g., Figures 4.13 and 5.13) with some discussion about how these scores should be used to help choose among options. In particular, given the wide confidence intervals for the scores, it seems important to discuss whether the outcomes for final VSP actually differ significantly among the scenarios.

7.B. Are existing data and scientific knowledge used effectively and appropriately?

This chapter seems like a useful place to summarize important data gaps and to prioritize research efforts for refining and verifying the modeling effort (as previously mentioned in our comments on Chapter 1).

7.C. Are assumptions and uncertainties about the analyses clearly described? For example, do the authors identify the strengths and weaknesses of the model and its inputs, and accuracy and precision of model output?

The conclusions are presented without adequately addressing uncertainty in the results that might be caused by some of the assumptions. The last chapter would benefit from a more explicit discussion of the caution that should be exercised in interpreting the model output.

A "No Hatchery" option is not considered for either basin. Why not? It may seem obvious that hatchery propagation is needed, but the model could help to confirm this presumption.

References

Backman, T.W.H., A.F. Evans, M.S. Robertson, and M.A. Hawbecker. 2002. Gas bubble trauma incidence in juvenile salmonids in the lower Columbia and Snake Rivers. North American Journal of Fisheries Management 22:965–972.

Caudill, C., M. Jepson, T. Clabough, S. Lee, T. Dick, M. Knoff, M. Morasch, M. Keefer and T. Friesen. 2013. Migration behavior and distribution of adult spring Chinook salmon radio-tagged at Willamette Falls. In 2012 Willamette Basin Fisheries Science Review. Oregon State University and the US Army Corps of Engineers, Corvallis, OR.

Chilcote, M., K. Goodson, and M. Falcy. 2011. Reduced recruitment performance in natural populations of anadromous salmonids associated with hatchery-reared fish. Canadian Journal of Fisheries and Aquatic Sciences. 68: 511-522.

Connor, W.P., H.L. Burge, and R. Waitt. 2002. Juvenile life history of wild fall Chinook salmon in the Snake and Clearwater Rivers. North American Journal of Fisheries Management 22:703-712.

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

Connor, W.P., C.E. Piston, and A.P. Garcia. 2003. Temperature during incubation as one factor affecting the distribution of Snake River Fall Chinook salmon spawning areas. Transactions of the American Fisheries Society 132:1236-1243.

Connor, W.P., S.G. Smith, T. Anderson, S.M. Bradbury, D.C. Burum, E.E. Hockersmith, M.L. Schuck, G.W. Mendel and R.M. Bugert. 2004. Post-release performance of hatchery yearling and subyearling fall Chinook salmon released into the Snake River. North American Journal of Fisheries Management 24:545-560.

Connor, W.P., J.G. Sneva, K.F. Tiffan, R.K. Steinhorst, and D. Ross. 2005. Two alternative juvenile life history types for fall Chinook salmon in the Snake River Basin. Transactions of the American Fisheries Society 134: 291-304.

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

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

Feist, B.E., E.R. Buhle, P. Arnold, J.W. Davis, and N.L. Scholz. 2011. Landscape ecotoxicology of Coho salmon spawner mortality in urban streams. *PLOS ONE* 6(8): 1-11.

Hamazaki, T., E. Kahler, B.M. Borba, and T. Burton. 2013. Impact of *Ichthyophonus* infection on spawning success of Yukon River Chinook salmon *Oncorhynchus tshawytscha*. 106: 207-215.

Hinch, S.G., S.J. Cooke, A.P. Farrell, K.M. Miller, M. Lapointe, and D.A. Patterson. 2012. Dead fish swimming: a review of research on the early migration and high premature mortality in adult Fraser River sockeye salmon *Oncorhynchus nerka*. J. Fish Biol. 81: 576-599.

Keefer, M.L., G.A. Taylor, D. F. Garletts, G.A. Gauthier, T.M. Pierce, and C.C. Caudill. 2010. Prespawn mortality in adult spring Chinook salmon outplanted above barrier dams. Ecology of Freshwater Fish 19:361-372.

McMichael, G.A., A.C. Hanson, R.A. Harnish and D.M. Trott. 2013. Juvenile salmonid migratory behavior at the mouth of the Columbia River and within the plume. Animal Biotelemetry 1:14.

Mesa M.G., L.K. Weiland, and A.G. Maule. 2000. Progression and severity of gas bubble trauma in juvenile salmonids. Transactions of the American Fisheries Society 129:174-185.

Parkhurst, Z.E., F.G. Bryant, and R.S. Nielson. 1950. Survey of the Columbia River and its tributaries. Part 3. U.S. Dept. of the Interior. Washington D.C. Special Scientific Report No. 36. 103 p.

Quinn, T.P., D.M. Eggers, J.H. Clark, and H.B. Rich, Jr. 2007. Density, climate, and the processes of prespawning mortality and egg retention in Pacific salmon (*Oncorhynchus* spp.). Canadian Journal of Fisheries and Aquatic Sciences, 2007 64: 574-582.

Romer, J.D., F.R. Monzyk, R. Emig, and T.A. Friesen. 2013a. Gill Na+K+ ATP-ase activity dynamic of juvenile Chinook salmon from McKenzie River and Cougar Reservoir. Addendum Report to: Juvenile salmonid outmigration monitoring at Willamette Valley project reservoirs. ODFW Corvallis Research Lab. 22 p.

Roumasset, A.G. 2012. Pre-spawning mortality of Upper Willamette River spring Chinook salmon associations with stream temperature, watershed attributes, and environmental conditions on the spawning grounds. M.S. Thesis College of Graduate Studies, University of Idaho. 118p.

Roumasset, A., and C. Caudill. 2013. Prespawn mortality of Upper Willamette River spring Chinook salmon: Associations with stream temperature, watershed attributes and environmental conditions on the spawning grounds. In 2012 Willamette Basin Fisheries Science Review. Oregon State University and the US Army Corps of Engineers, Corvallis, OR.

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

Schreck, C.B., M.L. Kent, M.E. Colvin, S. Benda, C. Sharpe, J.T. Peterson, and B. Dolan. 2013. Potential causes and management of prespawn mortality in adult Upper Willamette River spring Chinook. Annual Report. U.S. Army Corps of Engineers. Portland, OR. 62 p.

Tiffan, K.F., T.J. Kock, W.P. Connor, R.K. Steinhorst, and D.W. Rondorf. 2009. Behavioural thermoregulation by subyearling fall (autumn) Chinook salmon *Oncorhynchus tshawytscha* in a reservoir. Journal of Fish Biology 74:1562-1579.