Review of Methods for Forecasting Chinook Salmon Abundance in the Pacific Salmon Treaty Areas. Report to the Pacific Salmon Commission.

Randall M. Peterman
Ray Beamesderfer
Brian Bue

November 2016

Pacific Salmon Commission
Technical Report No. 35

The Pacific Salmon Commission is charged with the implementation of the Pacific Salmon Treaty, which was signed by Canada and the United States in 1985. The focus of the agreement are salmon stocks that originate in one country and are subject to interception by the other country. The objectives of the Treaty are to 1) conserve the five species of Pacific salmon in order to achieve optimum production, and 2) to divide the harvests so each country reaps the benefits of its investment in salmon management.

Technical Reports of the Pacific Salmon Commission present results of completed or ongoing investigations carried out by the Pacific Salmon Commission that are deemed of sufficient interest to be made available to the scientific community and the public.

Pacific Salmon Commission
600 - 1155 Robson Street
Vancouver, B.C.V6E 1B5
(604) 684-8081
www.psc.org

Pacific Salmon Commission
Technical Report No. 35

Review of Methods for Forecasting Chinook Salmon Abundance in the Pacific Salmon Treaty Areas. Report to the Pacific Salmon Commission.

Randall M. Peterman
Ray Beamesderfer
Brian Bue

November 2016

Correct citation for this publication:

Peterman, R.M., R. Beamesderfer, and B. Bue. 2016. Review of methods for forecasting Chinook salmon abundance : report to the Pacific Salmon Commission. Pacific Salmon Comm. Tech. Rep. No. 35: 165 p.

Acknowledgements

This report presents conclusions and recommendations of an independent panel charged with reviewing the methods for forecasting Chinook Salmon abundance for aggregate abundance-based management of fisheries under the U.S.-Canada Pacific Salmon Treaty. The Panel extends its gratitude to the many participants in the review workshop on forecasting methods in Portland, Oregon for their objective participation, critical perspectives, and patient answers to our innumerable questions. These participants included John Carlile and Bob Clark (ADFG); Antonio Velez-Espino, Diana Dobson, and Gayle Brown (CDFO); Stuart Ellis, Jon Hess, Marianne McClure, and Tommy Garrison (CRITFC); Alan Byrne (IDFG); Robert Kope (NMFS); Ethan Clemons, Jeff Whisler, Matt Falcy, Christine Mallette, John North and Tim Dalton (ODFW); Steve Haeseker (USFWS); Robin Ehlke, Ron Roler, Ben Cox and Lisa Harlan (WDFW); and Roger Dick, Jr. (YN Fisheries).

Table of Contents

1 EXECUTIVE SUMMARY

Management of Chinook Salmon harvest in the Pacific Salmon Treaty (PST) areas of Southeast Alaska and British Columbia under an Aggregate Abundance Based Management (AABM) framework depends heavily on forecasts of Chinook abundance prior to the onset of fishing (pre-season forecasts). In response to recent increases in the magnitude of differences between pre-season and post-season abundance estimates in the three AABM areas, as well as concerns about forecasts by regional agencies, the Pacific Salmon Commission (PSC) established an Independent Expert Panel. The mandate of the Panel was to (1) evaluate the accuracy and precision of stock-specific forecasts produced by agencies, and evaluate the differences between the PSC Chinook model's pre-season forecasts of abundance in AABM areas with post-season estimates, (2) provide advice on the strengths and weaknesses of forecasting methods and alternatives, and (3) suggest improvements. The Panel consisted of three members, one proposed by Canada and two by the U.S. Section.

An information-gathering workshop was facilitated by the Panel on 10-11 August 2016 in Portland, Oregon. Participants were responsible for agency forecasts of selected stocks or for forecasts from the PSC's Chinook model (the latter produced by the Chinook Technical Committee, CTC). A total of 23 non-panel participants attended, representing nine agencies. Presentations were made on agency forecasts and annual run reconstruction methods, and on the PSC Chinook Model's calibration and abundance-forecasting procedures. Before and after the workshop, the Panel reviewed large volumes of information from more than 70 documents and spreadsheets related to Chinook forecasting methods and results. Panel members also conducted extensive follow-up with forecasters and modelers by email and phone for additional information and explanations.

The Panel identified a number of issues affecting bias and precision of agency forecasting methods and their effective application – these are not necessarily weaknesses but are rather opportunities for improvement. For agency as well as PSC-model forecasting methods, the Panel identified general conclusions and specific recommendations for improvement. Issues, conclusions, and recommendations represent the consensus of the Panel members.

Recommendations are also qualitatively categorized by urgency and immediacy of potential implementation:

Near-term – Relatively straightforward to implement with likely immediate benefit (within 1 year).

Intermediate-term – Would require moderate investment of time and effort (1-2 years)

Long-term – Would likely require substantial time and effort, but with high potential for long term improvements (3-5 years).

Agency Forecasts

Bias & Precision

1. The forecast of Columbia Upriver Bright Chinook Salmon is nearly unbiased and reasonably precise (mean percent error = 1%; mean absolute percent error = 25%). The most recent returns (2013-15) were the largest in the data set and showed the greatest deviation between the forecast and actual terminal returns. There was no obvious time trend of forecasts being either under- or overestimates.
2. The forecast of Columbia River Spring Creek Hatchery shows a tendency to overestimate actual abundance and is reasonably precise (mean percent error = 8%; mean absolute percent error = 31%. There are periods of underestimates of the run (2001-2004) and overestimates (2005-2012; Figure 24).
3. The forecast of Columbia Upriver Summer Chinook Salmon was evaluated for the 2005-2015 returns. Overall, the forecast shows a tendency to overestimate abundance and is reasonably precise (mean percent error = 10%; mean absolute percent error = 22%). Forecasts for 7 of the 8 returns between 2005 and 2013 were greater than the observed returns, whereas the returns for 2014 and the record-high 2015 were more than forecast.
4. Forecasts of West Coast Vancouver Island Chinook Salmon are biased low (mean percent error = -27%) and are imprecise (mean absolute percent error = 42%). Fifteen of the seventeen forecasts examined were low with over forecasts occurring in 2005 and 2014.
5. Forecasts of North Oregon Coast Chinook Salmon tend to overestimate abundance but are reasonably precise (mean percent error = 8%; mean absolute percent error = 31%). However there was an unusually large over forecast in 2007 and when that forecast was removed from the analysis, there was a tendency to slightly underestimate the return (MPE = -2% and MAPE = 22%). Following Improvements in stock assessment in 2008 (more age sampling, increases in the speed of scale aging, and improved escapement estimation), forecasts have tended to underestimate the returning escapement (mean percent error = -6%) and forecast precision has been increased (mean absolute percent error = 14%).
6. A comparison of forecast bias (MPE) for the five stocks examined in this review with the forecast bias for 37 sockeye salmon and 40 chum salmon stocks examined in previous studies showed that the three Columbia River forecasts and the North Oregon Coastal forecast were at the low end of observed MPE values relative to sockeye and chum salmon forecasts. In contrast, the negative bias demonstrated by the West Coast Vancouver Island Chinook Salmon forecast (MPE) fell below the range of MPE observed for either chum or sockeye salmon.
7. A comparison of forecast precision (MAPE) for the five Chinook Salmon stocks examined in this review with the forecast precision for 37 sockeye salmon and 40 chum salmon stocks examined in previous studies showed that the three Columbia River forecasts and

the North Oregon Coastal forecast were at the low end of observed MAPE values for the sockeye and chum salmon forecasts. While the MAPE of 42% for the West Coast Vancouver Island forecast is higher than the other four Chinook Salmon stocks examined, it is still well within the range of MAPE values examined for chum and sockeye salmon.

Strengths

- ✓ Reflect the best available stock assessment data in any given year based on the resources and expert judgment of fishery biologists with the greatest familiarity with the stock.
- ✓ Forecasting methods are based on simple and relatively robust models with easily-understood methods and assumptions (given appropriate documentation).
- ✓ Provide generally comparable levels of bias and precision on average compared to those observed for other salmon species.
- ✓ Forecasts are generally useful for fishery management needs in both ocean and freshwater fisheries of British Columbia, Washington, and Oregon.

Issues

I. The current documentation of agency forecasts of abundance that are sent annually to the CTC does not provide sufficient information for PSC modelers to identify the long-term accuracy and precision of those forecasts, let alone uncertainty about the current year's forecast.

II. Efforts by agencies to provide forecasts as inputs to the PSC model are hampered by an incomplete understanding of (1) the PSC model's information requirements, (2) how those forecasts are used in that model, and (3) how those uses differ from those of fishery managers within regions.

III. The accuracy and precision of stock-specific forecasts are limited by the available stock assessment data; this is more of a problem for some Chinook stocks than others.

IV. There are substantial differences among regional agencies in how stock forecasts are produced and described. These differences cloud the interpretation of the point forecasts of abundance from the PSC model.

V. Forecasting methods for some stocks have not fully incorporated existing knowledge of changing parameters, such as age at maturity, or recent advancements in statistical methods of analysis.

VI. Existing forecasting models used by agencies, especially sibling relationships, are reasonably effective in representing average conditions but are vulnerable to performing poorly for years of very low or very high returns.

Suggestions for Improvement: Conclusions & Recommendations

A. <u>More comprehensive documentation is needed by the CTC from regional agency forecasters regarding the agencies' methods, critical assumptions and uncertainties, and</u>

accuracy and precision of past stock-specific forecasts. Agencies should also state the uncertainty in each stocks' annual forecasted abundance. More-frequent in-depth communication between PSC modelers and agency staff is also required.

1. *When regional agency forecasters send their stock-specific annual forecasts to the CTC, they should document their model-ranking procedures as well as the past performance of their methods (bias and precision). [See Recommendation 1 (Near-term)]*

2. *Agency forecasters should not choose just one best model for forecasting abundance in each age class. Instead, they should conduct analyses across different models that make different assumptions and report the resulting set of forecasts to the CTC for use as inputs to the PSC model. The generally large prediction intervals (not confidence intervals) around point forecasts should also be reported. [See Recommendation 2 (Intermediate-term)]*

3. *Agency forecasters should also send to the CTC a set of forecasts, each one based on a different model-ranking criterion, as determined by a range of management objectives. As described in section 8.1, the CTC can then conduct sensitivity analyses with the PSC model to determine their effect on forecasts of abundance in the AABMs. [See Recommendation 3 (Intermediate-term)]*

4. *All assumptions underlying the annual forecast, as well as data related to those assumptions, should be listed in the document provided to the PSC modelers so that everyone is aware of the forecast's strengths and weaknesses. [See Recommendation 17 (Near-term)]*

5. *A list of the alternative forecasting models examined and the criteria used to select among those models for producing a forecast for the Northern Oregon Coast should be clearly stated in the forecast document provided to the PSC model group, as suggested in recommendations at the start of section 7. [See Recommendation 16 (Near-term)]*

B. More explicit direction from the Chinook Technical Committee is needed by agency-based stock forecasters regarding the annually requested forecasts.

1. *The Columbia River Technical Advisory Committee and the Pacific Salmon Commission's modeling group should communicate with each other to ensure that they are both working with the same definition of the Columbia River Summer stock and the same sets of data, and that any historical information reflects this change. [See Recommendation 10 (Near-term)]*

2. *The CTC modeling group and WCVI (West Coast Vancouver Island) forecasters should decide (1) which type of forecast is required from WCVI (based on base-period data [1979-1982] or recent years, for example), and (2) the forecast*

performance values beyond which an extensive review of forecasting methods should be triggered. [See Recommendation 11 (Near-term)]

3. *The CTC should request each regional agency to provide to PSC modelers the forecasts of abundance for the model deemed best for each of the "relevant" ranking criteria (such as MRE, MAE, or RMSE), where "relevant" is defined as those that fit with stated management objectives for the AABMs . [See Recommendation 22 (Near-term)]*

C. Substantial improvements in basic assessments of some Chinook stocks are needed to support current PSC model and management applications, otherwise expectations need to be rescaled/reduced to recognize existing data limitations. Further expansion of the PSC model's number of stocks and fishing areas may need to be postponed until the quality of relevant data is deemed suitable.

1. Continue to improve upon the ability to estimate the contribution by stock to all AABM and ISBM fisheries with the objective of obtaining reliable stock contribution estimates by age. The Panel encourages the commitment of extra funding for the implementation of techniques to estimate stock contributions in a timely enough manner that the results can be used for forecasting in the subsequent year. *[See Recommendation 7 (Long term)]*

2. *An evaluation of the WCVI sampling program should be undertaken to determine if (1) there has been a dramatic change in sample collection methods and sampling intensity over the years, and (2) whether the sample design and intensity is adequate to obtain meaningful age composition estimates. If the sample design appears to be adequate, then explore other ways to estimate the age-3 and age-6 components of the returns. [See Recommendation 12 (Intermediate-term)]*

3. *Continue the increased sampling in the Northern Oregon Coast for age, rapid reading of scales for age, and improvements in escapement estimation. [See Recommendation 18 (Near-term)]*

D. Establishment of a set of “best forecasting practices" and standard definitions can improve the statistical foundation of methods for stock forecasting.

1. *We encourage all agency forecasters to apply ForecastR to their regions' stocks. As well, the CTC should run workshops to familiarize agency scientists with the ForecastR program. [See Recommendation 4 (Near-term)]*

2. *Explore the use of natural-log transformations for sibling regressions. The examination should evaluate both the effect on meeting the regression assumptions and forecasting performance. [See Recommendation 9 (Near-term)]*

E. Accuracy, precision, and transparency of stock forecasting methods might be substantially improved by application of more formal model-selection criteria that match clearly defined management objectives, as well as more advanced statistical methods that allow for time-varying parameters.

1. *The Columbia River Technical Advisory Committee (TAC) should explore whether using formal statistical model-selection criteria improves the accuracy and precision of their forecasts. [See Recommendation 8 (Near-term)]*

F. Development of new models and advanced parameter estimation methods may improve the accuracy and precision of agencies' annual forecasts. Regardless of any such improvements, large uncertainties in forecasts should be expected, especially when they are based on data outside the range of past observations.

1. *Agency forecasters should try applying a hybrid sibling model, especially to cases in which the fit of data to a standard sibling model is weak. [See Recommendation 5 (Near-term)]*
2. *We recommend that agency forecasters try using a Kalman filter estimation procedure for fitting their sibling relationships to account for time-varying parameters. [See Recommendation 6 (Near-term)]*
3. *The use of recent harvest rates and maturation rates should be explored for the WCVI forecasting model. These analyses should estimate model sensitivity to uncertainties in these rates. All results of these sensitivity analyses, including the associated forecasts, should be provided to CTC modelers along with estimates of uncertainty in the forecasts. [See Recommendation 13 (Intermediate-term)]*
4. *Explore a different and simpler method of forecasting terminal return to WCVI. The preferred method would reduce the complexity of the forecast by reducing the number of data manipulations and number of parameters and assumptions in the forecasting procedure. As with all new methods, it should be thoroughly evaluated to determine whether an increase in performance is actually obtained in terms of bias and precision. Sensitivity analyses should be performed to determine the influence of uncertainties in model parameters. [See Recommendation 14 (Intermediate-term)]*
5. *We recommend that Oregon Department of Fish and Wildlife (ODFW) forecasters examine log_e-log_e (natural logarithm sibling regressions, a hybrid sibling model, and a Kalman filter estimation procedure, the latter to account for possible temporal changes in parameters of the sibling relationship. [See Recommendation 15 (Near-term)]*
6. *As the population assessment models continue to evolve, North Oregon Coast (NOC) researchers should determine the sensitivity of the resulting forecasts to the uncertainty in estimated parameters in the models and quantify the uncertainty in the forecasts. [See Recommendation 19 (Intermediate-term)]*
7. *If more detailed data can be obtained from terminal fisheries for NOC, the forecast for this aggregate stock should change to a terminal run forecast instead of an escapement forecast. [See Recommendation 20 (Intermediate-term)]*

PSC model

Comparisons of pre-season forecasts and post-season estimates of AABM abundance[1]

The PSC model's pre-season forecasts of abundance in each of the three AABM areas can be evaluated by comparing them with the post-season abundance estimates in those same areas. However, both of these values are estimates from the PSC model. There is no independent estimate of actual abundances of Chinook in the AABMs. Thus, there is no way to calculate the bias and precision of the PSC model because there is no reliable way to estimate actual abundance of Chinook in the AABMs. This situation contrasts with approaches to evaluating stock-specific agency forecasting methods, in which catches and escapements can be summed to estimate actual returning abundances to compare with agency forecasts.

1. The magnitude of annual differences between the PSC model's pre- and post-season estimates of abundance in each of the three AABM areas have generally been less than 25% of the post-season estimates, but those deviations can represent hundreds of thousands of fish.
2. From about 2005 through 2011, forecasts tended to be greater than post-season estimates. For 2012 through 2015, forecasts in two years were overestimates and two were underestimates.
3. Large deviations between pre-and post-season abundance estimates from 2012 through 2015. Those deviations were highly positively correlated across the three AABM areas. That is, overestimates in Southeast Alaska occurred in years when overestimates occurred in Northern B.C. and the West Coast of Vancouver Island, and the same with underestimates.
4. Causes of the recent large discrepancies between the pre- and post-season AIs are unclear. However, the strong positive correlation in discrepancies across AABMs areas, along with other evidence, suggests that both the PSC model and the agencies' stock-specific forecasting methods do not properly represent changes in key factors such as time-varying maturation rates, marine survival rates, or exploitation rates.
5. Forecasts of Chinook Salmon obtained from the PSC model after the Agency forecasts were incorporated were relatively unbiased when measured by mean percent error (MPE) for four of the five stocks in this review. The forecast for the West Coast Vancouver Island stock was biased low (MPE=-17%) but not as biased as the Agency forecast (MPE=-30%).

---

[1] *A review of forecasts obtained from the PSC model absent input from the Agency forecasts for the five stocks in this review was not performed. Conversations with John Carlile (ADF&G) indicated that the model would need to be rerun with the Agency forecasts removed in order to determine how the PSC model would forecast absent Agency input. Given the large number of possible ways the model could be examined for the five stocks (one stock removed at a time, all stocks removed, or some combination), extremely limited staff time to do the model runs, and the scope of this review, it was determined that this evaluation would best be performed at a later date.*

The precision of the forecasts (MAPE) for the five stocks from the PSC model after the Agency forecasts were incorporated was comparable to that obtained from the Agency forecasts.

Strengths

- ✓ Combines stock-specific forecasts along with other data to produce forecasts of abundance of Chinook in the three AABMs. Those forecasts, which are in units relative to the base-period abundances (1979-1982) help determine maximum catches in AABMs based on the fishery control rules established by the Treaty.
- ✓ Extends terminal forecasts developed by the agencies to pre-fishery ocean abundance for application to AABM fisheries in Alaska and British Columbia.
- ✓ Calibration procedure incorporates current data into the forecasting method.
- ✓ Provides means of forecasting index and other stock abundances and returns to terminal fisheries.

Issues

VII. Incomplete and out-of-date documentation of the current PSC Chinook model and its calibration and projection procedures (1) threatens loss of institutional knowledge as key staff move on, (2) increases challenges to new CTC members who want to understand the model and its procedures, and in the worst case, (3) increases the chance of errors in the model's application and interpretation.

VIII. The deterministic nature of the PSC model and paucity of routine sensitivity analyses do not provide information about uncertainties in the model's forecasts of abundance in the three AABMs and terminal areas, thereby hampering well-informed decision making by PSC Commissioners and fishery managers in AABM areas.

IX. The PSC model's structure, parameterization, and calibration are complex and subject to substantial structural and parameter uncertainties.

X. Limitations of data and uncertainties associated with stock assessments and forecasting models challenge effective implementation of abundance-based management of Chinook under the Pacific Salmon Treaty.

Suggestions for Improvement: Conclusions & Recommendations

G. <u>Comprehensive up-to-date documentation of the PSC Chinook model in a single, central location is necessary to support its effective and credible use and improvement. A succession plan for training new model users is also critical.</u>

1. *Additional evaluation and documentation are needed of the PSC model's methods for dealing with stocks for which age-composition data and/or forecasts of terminal abundance or escapement are <u>not</u> available, given the large relative abundance of those stocks in some AABM areas. [See Recommendation 35 (Intermediate-term)]*

2. *The calibration procedure for the PSC model should be standardized and thoroughly documented to such an extent that a new member of the Analytical Working Group could repeat previous example analyses and come to the same stopping point about which calibration is deemed "final". [See Recommendation 39 (Intermediate-term)]*

H. <u>Point estimates of forecasts of abundance indices in the three AABM areas from the PSC model should be accompanied by descriptions of uncertainties in those forecasts. Uncertainties can be derived from extensive sensitivity analyses of effects of different assumptions and input parameters. Expression of uncertainty in these forecasts is essential for determining the confidence to be placed in them and allowing for appropriate consideration by fishery managers.</u>

1. *A series of projection runs should be conducted with the PSC model to produce a range of AIs for each AABM area. These AIs would reflect the different agencies' stock-specific model-ranking criteria that are deemed relevant to AABM management objectives. The set of projection runs will be reduced once the agencies clearly understand the AABM management objectives [See Recommendation 23 (Intermediate-term)].*

2. *Uncertainty in estimates from the PSC Chinook model should be explicitly represented either by making the model stochastic or running it across numerous sets of assumptions using sensitivity analyses. [See Recommendation 29 (Long-term)]*

I. <u>Substantial revision, testing, or possibly even replacement of the existing PSC Chinook model is necessary to effectively serve continuing needs, including the need for statements of uncertainty in the model's forecasts. A subgroup of CTC members should be created to explore such revisions and new models.</u>

1. *Functionality of the PSC Chinook model might be enhanced by including, where appropriate, nonlinear relationships such as those found in many other fisheries models, including the effect of fishing on reducing the fish abundance available to subsequent fisheries during a given year. [See Recommendation 24 (Intermediate-term)]*

2. *Effects of changes in marine spatial distribution of Chinook stocks on functionality of the PSC Chinook model need to be evaluated. [See Recommendation 25 (Intermediate-term)]*

3. *Sensitivity analyses with the PSC model should be used to explore different assumptions about (1) age structure for stocks without historical age composition data, (2) body-size structure used in the current method for estimating PNV, and (3) alternative structural formulations of the PSC model to calculate changes in age at maturity as a function of changes in body-size distributions. Some of those analyses could also assume various correlations with age-at-maturity schedules of other stocks. [See Recommendation 26 (Intermediate-term)]*

4. *The differences between pre-season and post-season abundance indices in each of the three AABMs might be reduced by including in the PSC model tendencies for multiple stocks to have positively correlated time series in productivities. [See Recommendation 27 (Long-term)]*

5. *The PSC model might be improved if factors such as EV and RT were calculated as functions of other variables. [See Recommendation 28 (Long-term)]*

6. *Ideally, the existing PSC Chinook model and/or its procedures should either be tested and refined or an entirely new model (or models) should be developed. [See Recommendation 30 (Long-term)]*

7. *Testing of the PSC model (and all other contemplated models) should be a high priority when the Data Generating Model is released. [See Recommendation 31 (Intermediate-term)]*

8. *Evaluations of the PSC model should include: (1) a check whether there is confounding of parameter estimates in the stage 1 calibration; (2) a series of sensitivity analyses/calibrations exploring alternative values for assumed age-specific natural mortality rates that might affect all other subsequent calculations and forecasts of abundance, and (3) consideration of whether the PSC model is being over-fit. [See Recommendation 32 (Near-term)]*

9. *Documentation should be provided regarding the basis of estimates of Ricker stock-recruitment parameters, as well as uncertainty in those estimates. Also, some improvement in performance of the PSC model might be obtained if the Analytical Working Group (AWG) used a Kalman filter that allows for a time-varying maximum productivity parameter in a given stock's Ricker stock-recruitment model. That Kalman filter procedure will explicitly take into account observation error as well as natural variation. [See Recommendation 33 (Intermediate-term)]*

10. *Given the large number of input parameters, all possible combinations of low, medium, and high values for each parameter may be impossibly time consuming. However, only a subset of those combinations would be needed to produce a range of forecast abundances. [See Recommendation 34 (Intermediate-term)]*

11. *The Panel generally recommends use of stock-specific forecasts provided by agencies rather than forecasts derived solely from the PSC model in the absence of clear evidence of improvements in accuracy and precision across multiple years for PSC model forecasts. [See Recommendation 36 (Near-term)]*

J. <u>Alternative forecasting frameworks, as well as ways of using forecasts of abundance, should be considered for Chinook if current information and resources are not sufficient to effectively conduct adequate analyses and implement provisions of the current Treaty. Those provisions may need to be changed during current negotiations.</u>

1. *Considerations of outcome uncertainty (deviation between desired and realized outcomes such as catches) , as well as uncertainties in forecasts, will influence expectations of managers of these AABM fisheries when they choose annual fishing regulations. [See Recommendation 37 (Long-term)]*
2. *The PSC Chinook model should take into account outcome uncertainty when making forecasts and presenting uncertainties in them. [See Recommendation 38 (Long-term)]*
3. *The abundance forecasts for AABMs areas produced by the PSC Chinook model should convey to managers the net effect of all of the major uncertainties described previously -- structural uncertainty, parametric uncertainty, uncertainty about management objectives, and outcome uncertainty. [See Recommendation 40 (Long-term)]*

2 INTRODUCTION

Chinook Salmon harvests in Pacific Salmon Treaty (PST) areas of Southeast Alaska and British Columbia are managed under an Aggregate Abundance Based Management (AABM) framework. AABM fisheries include southeast Alaska troll, net, and sport; northern British Columbia troll and sport; and West Coast Vancouver Island troll and sport. Annual maximum allowable landed catches in these three AABM fisheries are established based on an aggregate stock abundance index (AI) of contributing stocks to each AABM area. These AIs are calculated prior to each fishing season by the Chinook Technical Committee (CTC) using the Pacific Salmon Commission's (PSC's) Chinook model (henceforth "PSC model" or "CTC model"), based in part on stock-specific run forecasts derived by regional fishery management agencies and in part on numerous other sources of input data. [2]

For each AABM area, the Treaty identifies an abundance-based harvest control rule that allows for higher catches at greater aggregate abundances (Table 1 in Chapter 3 of the Treaty). There are several different AI tiers per fishery where the percentage harvest rate steps up to a higher level. Thus, higher abundance forecasts allow for higher harvests because more fish are available to the fishery, and above certain levels, the available fish are harvested at greater rates. Conversely, lower abundance forecasts require lower exploitation and produce lower harvests. There are also provisions in the Treaty that reduce AI catch levels when selected stock and stock aggregates are below conservation objectives recognized by the PSC.

During the recent period of widely variable Chinook abundance throughout the north-eastern Pacific, differences between pre-season (i.e., pre-fishery) abundance estimates in the AABMs and their respective post-season abundance estimates have increased considerably. Relatively large deviations have also been observed for specific stocks between recent pre-season forecasts of terminal run sizes (or escapement) and post-season estimates. Thus, various concerns have been raised about pre-season forecasts provided by agencies as input to the annual calibration procedure of the PSC coast-wide Chinook model (CTC model). To address these concerns, the PSC approved a process and timeline for an independent technical review of "... three methods for predicting stock abundance (agency forecast, CTC model calibration from agency forecast, and CTC model forecast absent agency forecast)". An Independent Technical Panel ("the Panel") was established to do this review.

For agency forecasts, the Panel was asked to focus on five Chinook stocks that have substantial abundance in AABM areas and/or have had recent forecasting performance issues:

- Columbia River Upriver Brights,

[2] *In this document "agency" will refer to the Alaska Department of Fish and Game (ADFG), Fisheries and Oceans Canada (CDFO), Washington Department of Fish and Wildlife (WDFW), Oregon Department of Fish and Wildlife (ODFW), Columbia River Inter-Tribal Fish Commission (CRITFC) or the U.S. Fish and Wildlife Service (USFWS).*

- Columbia River Summers,
- Spring Creek (Columbia River),
- Northern Oregon Coastal Fall, and
- West Coast Vancouver Island.

Objectives of the review include, but are not limited to:

1) Evaluate the bias and precision of alternative methods for predicting the pre- and post-season abundance (Abundance Index, AI);
2) Provide advice on the strengths and weaknesses of each method; and
3) Suggest improvements to current agency pre-season forecast methods for predicting stock abundance.

Additional details regarding this review's objectives and process may be found in the PSC's executive summary of the "Terms of Reference" (see Appendix A).

3 THE PANEL'S PROCESS

After approval of this process by the Pacific Salmon Commission, an independent panel of scientists was appointed from nominations by the respective delegations. The Panel consisted of three members, one proposed by Canada (Randall Peterman) and two by the U.S. Section (Brian Bue and Ray Beamesderfer). Brief biographies of Panel members may be found in Appendix B. The PSC Secretariat also alerted agencies affected by the review process and requested that pertinent information be provided to the Panel for review.

An information-gathering workshop was held by the Panel, CTC, and Agency representatives on 10-11 August 2016 at the Columbia River Inter-Tribal Fish Commission in Portland, Oregon (see Appendix C for details of workshop agenda, participants, and PowerPoint presentations). Participants included people directly responsible for the selected agency forecasts and for forecasts from the PSC Chinook model. At the workshop, agency staff provided technical/analytical presentations about the domestic agency forecasts and annual run reconstruction methods. Presentations were also made about the PSC Chinook model in order to familiarize the Panel members with its structure and calibration procedures for incorporating the agency-provided forecasts, as well as other data, including where no agency forecast was available, and forecasting pre-fishery abundances for allocation to mixed-stock ocean fisheries.

The workshop was facilitated by the Panel and encouraged discussions regarding:

- Problems or issues affecting bias, precision, or use of Chinook abundance forecasts, and
- Central issues or improvements in forecasting methods identified by workshop participants for consideration by the Panel with respect to both stock-specific forecasts and the PSC Chinook model's forecasts (see Tables 2, 3, and 4 in Appendix C).

Following the workshop, the Panel reviewed large volumes of information related to Chinook forecasting methods and results. Extensive follow-up was conducted by the Panel with forecasters and modelers by e-mail and phone to obtain additional information and explanations. The Panel conferred frequently by conference calls and e-mail to discuss information, identify key and issues, draw conclusions, and develop recommendations. All conclusions and recommendations in this report represent the consensus of the Panel members.

A draft report was provided to the PSC on September 15, 2016 and distributed to participating agencies for review and comment. Review comments were provided by the agencies with a deadline for receipt of October 1, 2016. Randall Peterman presented preliminary findings to the PSC on October 6, 2016. The PSC also provided additional suggestions on October 24, 2016 regarding the structure of the final report, particularly in the executive summary. The Panel subsequently revised the draft report to address comments and suggestions by the reviewers.

Descriptions of general issues and conclusions regarding forecasting methods are found in Section 6 of this report. More detailed recommendations are found in Section 7 for forecasting models used by Agencies for the five stocks reviewed and in Section 8 for the PSC Chinook model.

4 BACKGROUND

This section summarizes Portland workshop's presentations regarding the methods used for making agencies' stock-specific forecasts and PSC model forecasts. The Panel was asked to describe these methods in this report. Thumbnails of PowerPoint presentations from the workshop may be found in Appendix C. The summaries below include some quotes or paraphrases of those slides plus related points from reports or other material provided to the Panel. More detailed descriptions and the Panel's analyses appear in sections 7 and 8.

4.1 Pacific Salmon Treaty and Fisheries

Gayle Brown (CDFO) provided a brief history of the Chinook management framework under the Pacific Salmon Treaty and the key tool supporting this framework – the PSC Coast-Wide Chinook Model (the PSC model, also called the CTC model).

The 1985 US-Canada Pacific Salmon Treaty addresses all species of Pacific salmon and fisheries from Cape Falcon, Oregon to Cape Suckling, Alaska. Chinook stocks were generally believed to be depressed coast-wide by the late 1970s and early 1980s. The Treaty was a recognition that a Chinook rebuilding program was required. A rebuilding assessment tool was needed as well.

A Chinook Technical Team (CTC) was appointed under the PST. This team reports to the PSC and includes tribal and agency representatives from Alaska, British Columbia, Washington and

Oregon. Responsibilities include production of extensive yearly reports on Chinook catches and escapements, coded wire tags (CWT), exploitation rate analysis (ERA), and PSC Chinook Model calibrations. Other analyses and reports are also prepared at the request of PSC Commissioners. The Analytical Work Group (AWG), a subgroup of the CTC, produces annual ERA reports and conducts the PSC Chinook Model calibrations and projections of abundance.

The PSC Chinook Model is a cohort analysis model used for assessment. The first version was developed in 1983 with one stock. The model was transferred to BASIC computer code in 1984 and was subsequently ported into newer versions of BASIC code over the years, including the most recent Visual Basic version. In 1985, it was expanded to five stocks and ten fisheries, and those numbers have periodically increased over the years. In 2010, the model was expanded to include 30 stocks and 25 fisheries. Current plans are to expand the PSC model to 40 stocks and 48 fisheries in the future. Initially, the model was used for evaluation of management strategies (catch ceilings, harvest rates, etc.) as the basis for a 15-year rebuilding program. Since 1999, the model has been used as a management tool by providing forecasts of Chinook abundance for the three AABMs, as described in the Introduction.

4.2 Pacific Salmon Commission's Chinook Model

John Carlile (ADFG) and Antonio Velez-Espino (CDFO) described the PSC Chinook model in greater detail. The model is deterministic with annual time periods. All fisheries act on a single pool of fish (no explicit migration occurs among fisheries). Data are incorporated from CWT-based cohort analyses, as are historical data on catch and terminal run/escapement. Abundance is scaled to exploitation rates from a base period (1979-1982).

The model assumes that the ocean distributions of individual stocks are the same as those experienced during the model base period, i.e., static. Hatchery indicator stocks are treated as surrogates for wild stocks in the same geographic area with similar life histories (i.e., age structure, maturation rate, ocean distribution). All stocks of a given age have the same size distribution in a given fishery.

The Pacific Salmon Treaty dictates that management of Chinook fisheries in the three AABM areas is tied to pre-season estimates of abundance indices produced by the PSC Chinook model, so there has been a reluctance to modify or replace that model.

Inputs to the PSC model come from 11 input files that include base-period CWT data, fishery catch data, Chinook non-retention data, past escapement and/or terminal run data, terminal run/escapement forecasts, fishery policy (FP)-exploitation-rate scalars, maturation rate and adult equivalent data, hatchery enhancement data, spawner-recruit parameters, changes in proportion of fish not vulnerable to fishing gear (PNV), and inter-dam loss factors.

Calculations include two calibrations followed by a projection run to generate abundance forecasts for the three AABMs. A CWT Recovery Program summarizes base-period CWT data by stock. A Base Calibration Program consisting of a backward cohort analysis computes base-period exploitation rates, initial cohort abundances, and spawner-recruit parameters. The PSC Chinook model then runs a forward cohort analysis that fits to catches, escapement, terminal runs, agency forecasts, and other data.

Exploitation-rate scalars (FP) account for changes in time and area openings and changes in size limits that have occurred after the base period, 1979-1982. Between 1986 and 1998, 17 of the 25 PSC fisheries had catch ceilings. The RT scalar is used by the Chinook Model to adjust the legal catch to match the observed catch under ceiling management. Thus, stage-1 calibration estimates the RT scalars for ceiling fisheries in order to reproduce the observed ceiling catches. Stage-2 calibration fine tunes the EV scalar estimates that adjust the base-period smolt-to-age 1 survival rates such that the observed escapements and terminal runs are reproduced by the model. The projection run then produces the pre-season estimates of cohort abundance indices for each AABM.

The model calculates the terminal runs (cohort size minus ocean harvest), escapement (terminal run minus terminal harvest), and age 1 cohorts for the next year (from escapement fed into a spawner-recruit function). A starting cohort size is supplied for the first year only, along with average natural mortality rates, average base-period harvest rates, average maturity rates, and average spawner-recruit parameters. Inputs also include observed catches, escapements, and terminal runs. The model loops through all years in the database starting in 1979, estimating the cohort abundances up to the current year.

EV factors scale the number of recruits produced by the stock-specific spawner-recruit parameters to match supplied escapement/terminal run values. EV factors are stock- and brood-year specific, so abundance by age is used in the model. Age compositions are either fed into the model from observed data or generated based on maturity rates from the base period, 1979-1982. Stage-2 calibration generates the EVs that are to be used for the 1-year-ahead projection run, which produces the forecasts of abundance. EVs for different stocks are interrelated (i.e., each iteration of the calibration will potentially change the EVs for all stocks and brood years). If the spawner-recruit parameters are appropriate for a stock, then the EVs can be thought of as survival scalars. If they are not appropriate, then the EVs can be thought of as survival scalars combined with other factors that are assumed constant in the model but that in fact vary in nature.

Agency forecasts for specific stocks are used as inputs to the PSC model's calculation of annual abundance indices for the three AABMs prior to the next fishing season. These stock-specific agency forecasts have generally been used in annual calibrations of the PSC model without being scrutinized by the CTC. Model inputs for the 2016 forecasts include 28 stock-specific sets of time series data, 9 for escapement and 19 for terminal run. All input time series include observed (actual) data starting in 1979. Of the 28 stocks' input time series, 22 include agency-derived forecasts of abundance for that stock, 16 include historical age-composition data, but only 12 stocks have age-specific forecasts of abundance. Agency forecasts are prepared using the wide range of methods as described later. Model stocks range from individual stocks (Nooksack Springs) to large aggregates of many stocks (Fraser Early). Aggregates are usually mixtures of natural spawning stocks and hatchery stocks. Forecasts are needed in March by the PSC modelers when data from the previous year may not be available because of delays in processing coded-wire tags, but regardless, stock-specific forecasts are produced.

The PSC Chinook model generates stock-specific forecasts of the terminal run (or escapement, depending on the stock) regardless of whether an agency forecast is provided. Where no agency forecast is available, the PSC model's forecast is produced using recent averages of EV, along with many other assumptions. The model calibrates (fits) to the total brood-year terminal run/escapement, but uncertainties exist in estimates of age composition (maturation rate), exploitation rate, etc., especially in stocks for which only total abundance (no age structure data) is provided as input to the PSC model.

Effects of stock forecasts on AABM fishery AIs are related to the proportion of total abundance in a given AABM area that each stock represents, with major contributors having more influence on the AI. Such effects also depend on the magnitude of differences in pre-fishery cohort sizes among stocks, which are in turn affected by the PSC model's stock-specific EVs produced with and without agency forecast data. Accuracy (and age composition) of agency forecasts for all stocks have an effect on AIs, in part through their influence on estimates of recent EVs. Interactions with other input data and assumptions (e.g., FPs, maturation rates, etc.) also have an effect on AIs.

Outputs from the PSC model include (1) catches by fishery, stock, and age, (2) incidental mortalities by fishery, stock, and age, (3) fishery-specific stock composition estimates, (4) exploitation rates by fishery, stock, and age, (5) terminal runs/escapements by stock and age (original intent of the model), and (6) abundance indices (AIs) for the Southeast Alaska, Northern British Columbia, and West Coast of Vancouver Island AABMs (current focus of the model). The area-specific abundance index, AI, is calculated as a ratio -- the model's forecasted catch in each fishery (assuming 1979-82 base-period exploitation rates and current-year abundances) divided by the catch under base-period exploitation rates and base-period abundances.

The pre-2013 PSC model applied long-term average maturation rates to recent incomplete broods when calculating AIs. However, it was discovered that a number of stocks were maturing at younger ages than in the past (CTC 2016a). Such younger fish are in reality less vulnerable to fishing gear than older fish, but the model was assuming an unchanged, more vulnerable historical older-aged structure. Hence, because AIs are calculated for vulnerable cohorts, the PSC model's estimated abundance was too high for affected stocks. A fix was implemented in the 2013 model calibration by replacing long-term average maturation rates with the recent 5-year-average. In 2016, the average maturation rate that had been applied to incomplete cohorts was re-examined again, and it was found that the recent 9-year average performed better than the 5-year average.

This 2013 change in assumed maturation rates coincided with an end to the chronic over-prediction of the pre-season AIs, but it is still too early to determine whether the source of that bias has been removed. In 2013 and 2015, pre-season forecasts of AIs underestimated post-season AIs in all three AABMs, and in contrast, 2012 and 2014 pre-season forecasts of AIs were too high (Figure 1).

Causes of the recent large discrepancies between the pre- and post-season AIs are unclear. It is unknown how much is due to inaccurate terminal run/escapement forecasts provided by the agencies, as opposed to other sources. As a result of these questions, no CTC consensus was reached on 2015 or 2016 Model calibrations – instead, decisions were settled by the Commission. These concerns led to establishment of this review process.

There is another key point about the deviations between pre- and post-season AIs in Figure 1. A previous CTC analysis investigated the association between annual discrepancies in agencies' forecasts and the PSC model's deviations between pre- and post-season estimates in the AIs for the AABMs (CTC 2014). The combined error of stocks with the largest contributions (> 5%) to AABM-specific AIs is highly positively correlated with errors in PSC model's forecasts of AIs in the SEAK, NBC, and WCVI AABMs (r^2 = 0.7, 0.6, and 0.55, respectively) (p. 126 of CTC 2014). Of course, this is just a correlation and does not necessarily reflect a causal relationship between agency and PSC model forecasting discrepancies. Another possible reason for the high correlation (not mentioned by the CTC) is that both types of forecasting methods do not properly represent changes in key factors such as time-varying maturation rates, marine survival rates, or exploitation rates, thereby producing similar errors in particular years.

Some support for the latter interpretation is provided by the high positive correlation across the three AABM areas in their annual deviations between pre-and post-season abundance estimates, starting in 2012 (Figure 1). The stock composition differs considerably across the three AABMs, so it is unlikely that errors in stock-specific forecasts would explain that positive correlation in AI discrepancies across AABMs. A more likely explanation is that there have been major changes in large-spatial-scale factors such as maturation rates (which reflect growth rates) or marine survival rates that neither the regional forecasting models nor the PSC Chinook model have fully accounted for. The implication is that both of these types of models should be improved by explicitly estimating and using these time-varying parameters in their forecasting.

The presenters at the Portland workshop identified some alternatives to the PSC Chinook model. A model using continuous catch equations was proposed in 2004 by Gary Morishima and Din-Geng Chen, funded by the US LOA (Letter of Agreement). This model could potentially better account for interactions between fisheries, make temporal stratification of fisheries easier, and provide more information on the variability of stock distributions. However, we were told that the disadvantage of this model is that estimates of effort would be needed for each fishery, and that those estimates are not readily available for some fisheries. Another alternative model is a catch-at-age model proposed in 2004 by Rishi Sharma and Henry Yuen, but it was never followed up.

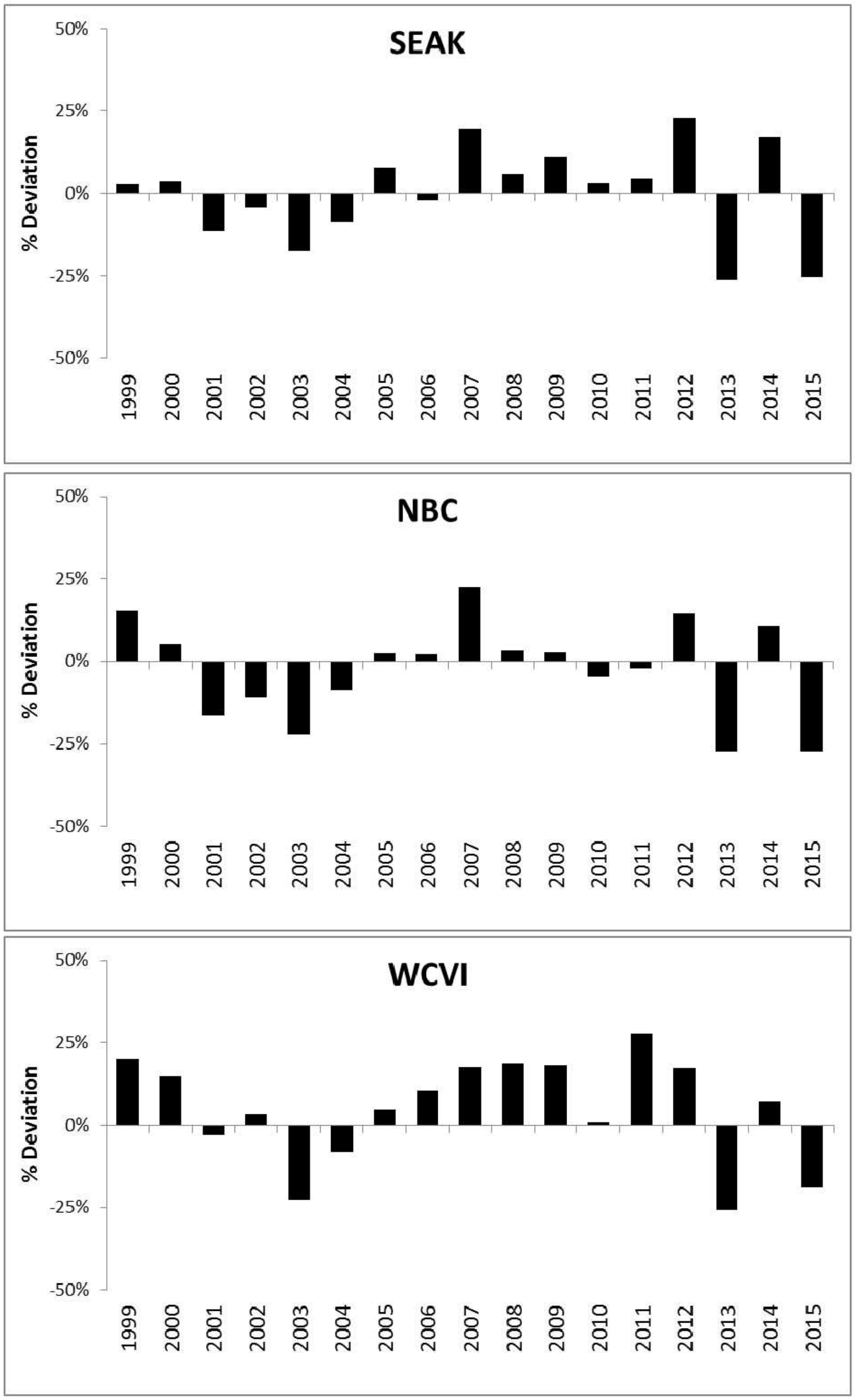


Figure 1. Deviations between pre- and post-season Chinook Salmon abundance indices, ([pre-season forecast - post-season]/post-season)*100, derived from the PSC model for the three AABM fisheries (CTC 2016b). **A positive deviation means the pre-season forecast was too high.**

A Data Generation Model (DGM) is currently being developed that will allow sample datasets to be generated, which will allow for comparison of output statistics from different forecasting models against known parameters (cohort sizes, exploitation rates, etc.).

4.3 West Coast Vancouver Island Forecast

Diana Dobson (CDFO) reported that the WCVI terminal run forecast includes Chinook from three major hatcheries and 18 index stocks, many of which are enhanced. This terminal index likely accounts for greater than 95% of annual WCVI Chinook production, which averages about 150,000 fish (range 40,000 to 300,000) (Figure 2). The average terminal age composition is 30% age 3s, 50% age 4s, and 20% age 5s. Substantial harvest of WCVI Chinook occurs in AABM fisheries of Southeast Alaska and Northern British Columbia, as well as in terminal fisheries of WCVI (Figure 3).

The forecasting method for WCVI is a complex, multi-stage process. It begins with linear "sibling" regression models (abundances are on the arithmetic scale, not logged) to predict the production (abundance) of older age classes from the observed production of younger age classes from the same brood year. Regressions are developed for CWT-associated production from the Robertson Creek Hatchery (RCH) Indicator Stock, which is then expanded to the entire Somass/RCH system. The terminal return of Somass/RCH is then predicted after applying assumptions about pre-terminal fishing mortality, stock composition in pre-terminal fisheries, and maturation rates. The forecast terminal return of Somass/RCH is then expanded for the WCVI index systems.

More specifically, for the Robertson Creek Hatchery (RCH) CWT Indicator Stock, simple linear sibling regressions are developed from production data generated by a cohort analysis, which is based on its own assumed natural morality and maturation rates. Two sibling regressions are computed, as described in more detail in section 7. Model Prod2 uses <u>total terminal return</u> at a younger age class (independent variable) to predict the dependent variable, total production (the surviving cohort in the ocean, i.e., ocean fishing mortality plus terminal run) of a subsequent age or ages from the same brood year. In contrast, Model Prod3 uses estimated <u>total production</u> (total fishing mortality plus escapement) <u>of a younger age class(es)</u> to predict total production of subsequent ages from the same brood year (again, the surviving cohort in the ocean). The forecast for the CWT-associated production for Robertson Creek Hatchery is then expanded for the entire Somass/RCH system based on ratios of earlier returns from the brood year. After the Somass/RCH production for each brood year is forecast, some assumptions are then applied to predict the terminal run size of Somass/RCH Chinook. Those assumptions include an assumed pre-terminal fishing mortality, pre-terminal fishery stock composition, and maturity rate. The latter two are generated by the cohort analysis.

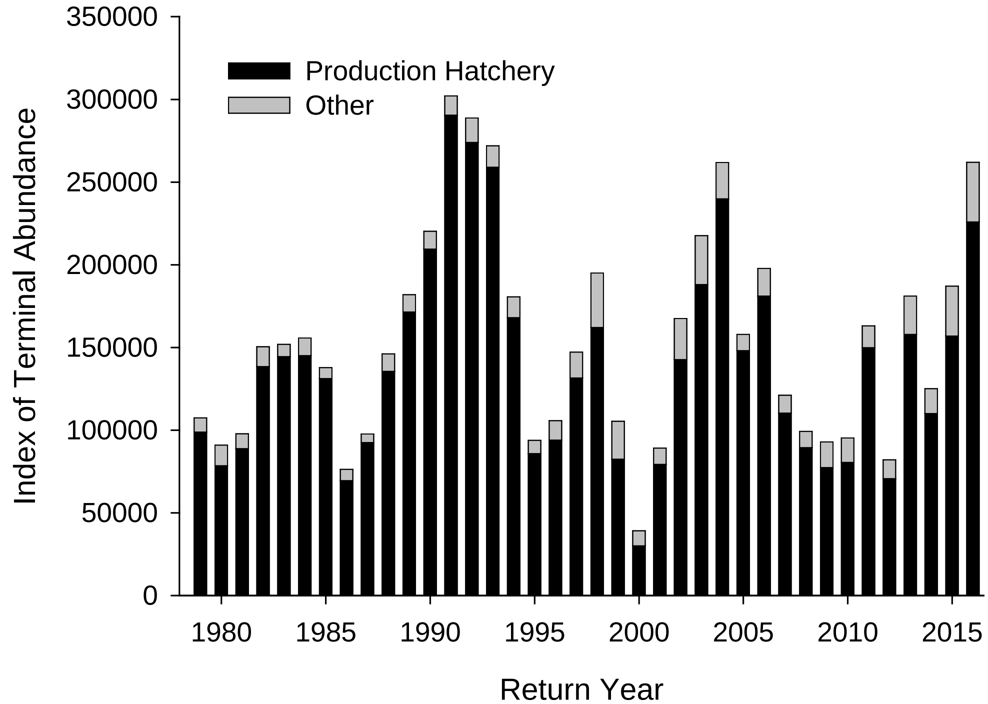


Figure 2. Index of terminal abundance for West Coast Vancouver Island Chinook Salmon (Diana Dobson, CDFO, workshop presentation).

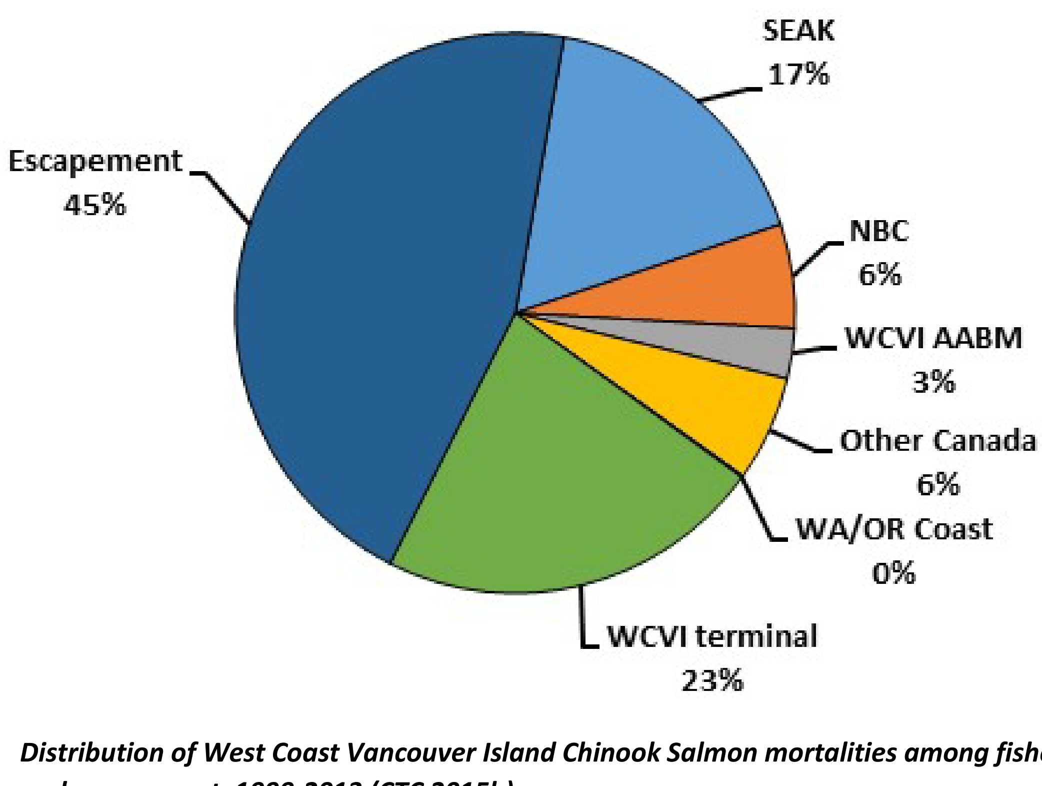


Figure 3. Distribution of West Coast Vancouver Island Chinook Salmon mortalities among fisheries and escapement, 1999-2013 (CTC 2015b).

Pre-terminal fishery mortality needs to be predicted for each age class/brood year. There is some uncertainty as to the correct assumption for WCVI forecasts to be used for PSC model calibration purposes. Is it the pre-terminal fishery mortality that was exerted during the base period, or is it the expected pre-terminal fishery mortality given current fishing regimes? In recent years, both options have been presented, although the latter has been used as the input of WCVI terminal run to the PSC model for its calibration.

The Somass/RCH terminal forecast was expanded for the WCVI index stocks with a similar method to that used for the expansion from RCH CWT production to Somass/RCH total production. In recent years, the Somass/RCH terminal forecast has also been expanded for the WCVI index stocks by adding terminal forecasts that are generated separately for Conuma hatchery and Nitinat hatchery returns and the 18 other index stocks combined. The forecasts generated for those other stocks use information from the RCH CWT cohort analysis (i.e., estimated brood-year survival rate) and similar pre-terminal fishery assumptions, but are modified with stock-specific production and age data.

All years of RCH CWT cohort data are incorporated in the sibling regressions that form the basis of the WCVI forecast (brood year 1983 and onward). Similarly, the WCVI terminal run index has been reconstructed from return year 1979 onward. All data are incorporated into the forecast and analysis. The more challenging issues relate to the varying quality of available assessment data across WCVI systems. There is a general paucity of data for WCVI stocks other than the RCH CWT indicator stock and, in some WCVI areas, few sample data from fisheries. There are low recovery rates for age 2 fish from which age-3 of the same brood are estimated, and there is a known bias in CWT-recovery data where individual stocks do not comprise a substantial percentage of the catch. A key assumption for WCVI forecasting is that the RCH hatchery indicator stock has similar survival rates, maturation rates, spatial distribution, and exploitation pattern to those of wild WCVI stocks.

In recent years forecasts of WCVI terminal run abundance have consistently underestimated actual values calculated after the fishing season by the PSC Chinook model (see section 7).

Suggestions for improvement in forecasts by Diana Dobson included:

- Resolve what input is required for calibration purposes – build a common understanding.
- Succession and documentation requirement for the process in general – misunderstanding or miscommunication of objectives and/or structural modifications could be a source of error.
- Age 3 forecast – a clear structural issue, also an input problem (age 2 data), and also likely related to changing maturation rate – could explore ‘leading indicators’ as an alternative method for Age 3 forecasting.
- Input data; not all available information is being used (e.g., available mark data and technology such as DNA, otolith marks, etc.)

- Incorporation of uncertainty – not just adding uncertainties to reports of input forecasts, but incorporating them into the entire management and assessment framework for Chinook.
- Simplification of the assessment and management framework – currently data intensive, assumption laden, and staff-limited.
- Separation of hatchery from wild abundance in the AIs.

4.4 Columbia River Upriver Bright Forecast

Jeff Whisler (ODFW) and Steve Haeseker (USFWS) reported that most upriver bright (URB) fall Chinook Salmon are naturally-produced and destined for the Hanford Reach area of the Columbia River. This stock also returns to the Priest Rapids Hatchery, areas upstream of Priest Rapids Dam, the Snake River, the Deschutes River, and the Yakima River. During 1980-2015, the mean return to the Columbia River was 246,300. Although there is year-to-year variability, on average 25% of fish return at age 2, 20% at age 3, 37% at age 4, 17% at age 5, and 1% at age 6 (Jeff Whisler, ODFW, personal communication).

CTC reports of the PSC Chinook model outputs estimate the spatial distribution of harvest of this stock. Substantial harvest occurs in the AABM fisheries of Southeast Alaska and Northern British Columbia, as well as terminal fisheries of the Columbia River (Figure 4).

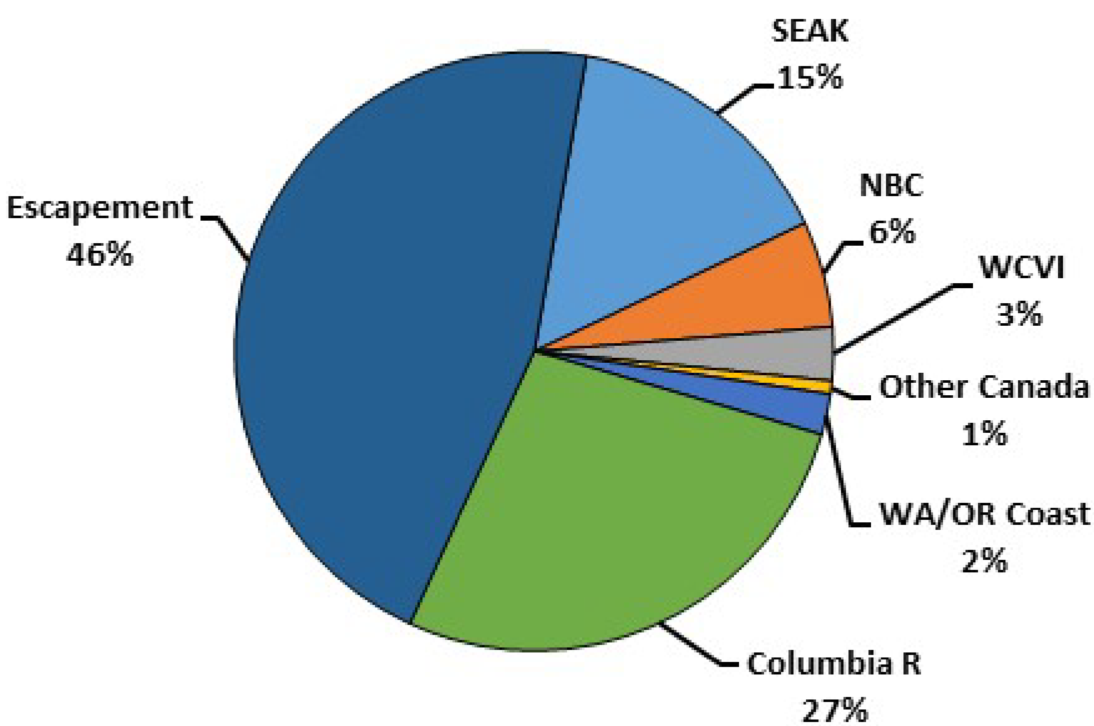


Figure 4. Distribution of Columbia Upriver Bright Chinook Salmon mortalities among fisheries and escapement, 1999-2013 (CTC 2015b).

Annual forecasts of Columbia River upriver bright fall Chinook returns are produced by an expert panel that includes members from WDFW and the *U.S. v. Oregon* Technical Advisory Committee (TAC). The quality of the data on total returns to the Columbia River and the age-composition of those returns is relatively high due to extensive sampling of Columbia River fisheries and recoveries of coded-wire tags (CWT). The primary forecasting methods are arithmetic-scale sibling regressions and average cohort ratios (e.g., the average ratio of age-3 to age-4 returns from the same brood year). When sibling regressions are explored with input data from different periods, the r^2 value is used to select the best model (Table 1 of WDFW 2016). The approach used to produce forecasts for Columbia River fall Chinook stocks is a modified Delphi method (i.e., open discussion among the expert panel). Output of a suite of stock- and age-specific models is presented to the panel and the merits of each are discussed before the panel comes to consensus. In the past few years, when returns have been setting modern-day record highs, the panel has relied on cohort ratios when regression inputs have been outside the range of the dataset.

In his workshop presentation, Steve Haeseker of the USFWS reported that natural variability in age composition makes forecasting of upriver brights difficult, but recent forecasts have been relatively precise and unbiased. For 1980-2015, mean percent error (MPE, a measure of long-term statistical bias) was -5% and mean absolute percent error (MAPE, a measure of precision) was 20% (Figure 5). However, substantial under- or over-estimates (up to about $\pm$50%) of forecasted abundance have occurred occasionally since 2001.

Figure 6 provides an example of a sibling relation showing (1) the typical very large 95% prediction interval, which illustrates the wide range across which age-3 abundances are likely to occur (with a probability of 95%) for a given age-2 abundance from the same brood-year cohort, and (2) the effect of between-year changes in maturity rate and/or survival rates. Both issues create large challenges for forecasting.

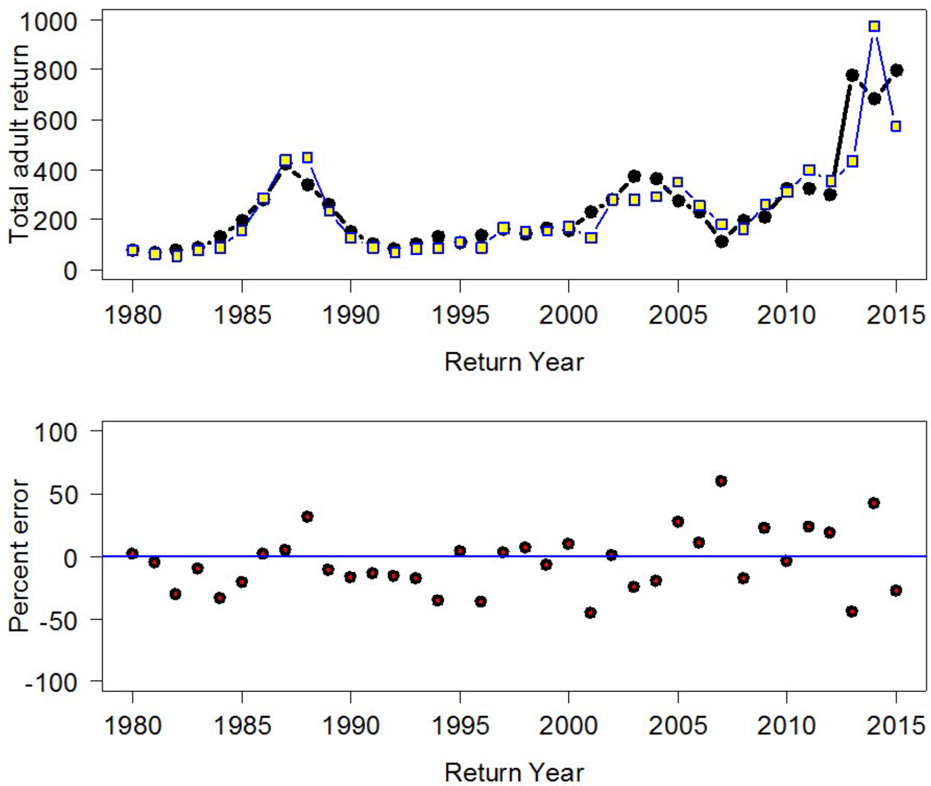


Figure 5. ***For return years 1980-2015, post-season estimates of abundance (thousands, black dots) and pre-season forecasts (yellow squares) (top panel), and percent error, ([forecast – actual]/actual) *100, between the pre-season forecasts and post-season abundance estimates for Columbia River mouth returns of Columbia Upriver Bright Chinook Salmon (Steve Haeseker, workshop presentation). Positive errors mean the forecast was higher than the actual return.***

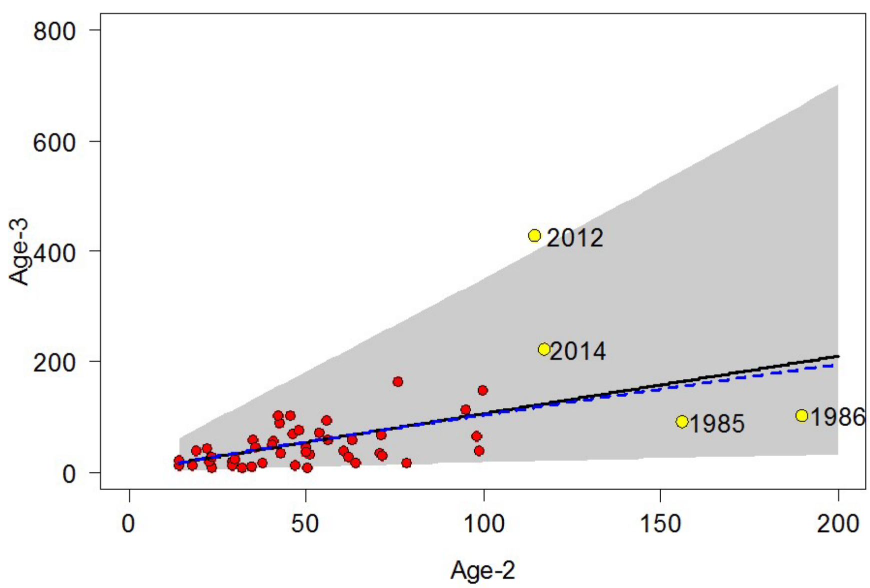


Figure 6. ***Example of a sibling relationship for Columbia River upriver bright Chinook abundances (in thousands), brood years 1962-2012, showing the 95% <u>prediction</u> interval (Steve Haeseker, workshop presentation). Years with extremely high age-2 abundances have yellow dots.***

4.5 Columbia River Spring Creek Hatchery Forecast

Steve Haeseker (USFWS) reported that this stock is produced by the U.S. Fish and Wildlife Service's Spring Creek National Fish Hatchery (NFH), located 35 km upriver of Bonneville Dam. The hatchery currently produces 10.5 million sub-yearling (ocean-type) tule fall Chinook Salmon annually. Tule fall Chinook Salmon are native to this part of the Columbia River and originally spawned in the White Salmon River one mile east of the hatchery. During 1980-2015, the mean return of Spring Creek tule fall Chinook to the Columbia River was 65,700. Although there is year-to-year variability, on average 8% return at age-2, 60% at age-3, 30% at age-4, and 2% at age-5 (Jeff Whisler, ODFW, personal communication).

CTC reports of PSC Chinook model outputs estimate that substantial harvest of this stock occurs in the West Coast Vancouver Island AABM fishery and in ISBM fisheries of the Washington/Oregon Coast to the Columbia River (Figure 7).

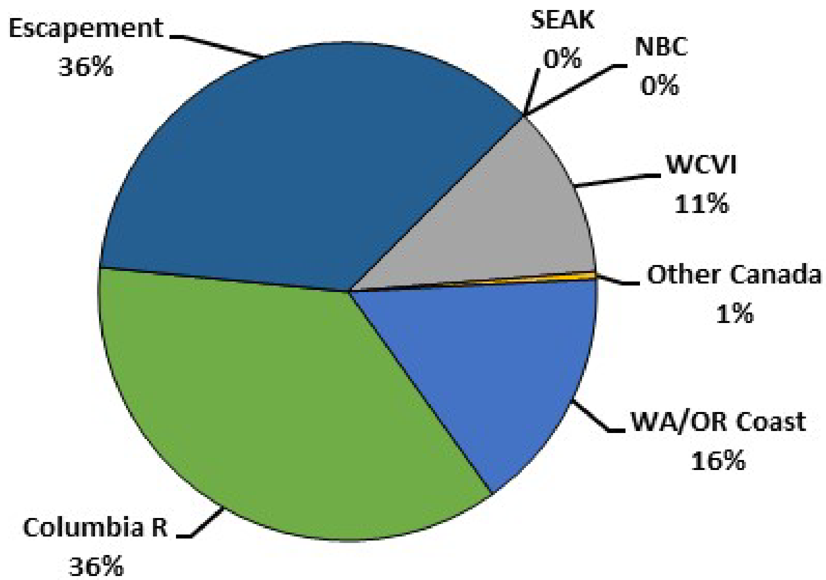


Figure 7. Distribution of Columbia River Spring Creek Hatchery Chinook Salmon mortalities among fisheries and escapement, 1999-2013 (CTC 2015b).

Annual forecasts of Columbia River Spring Creek Hatchery fall Chinook returns are produced by an expert panel that includes members from WDFW and the U.S. v. Oregon Technical Advisory Committee (TAC). The quality of the data on total returns to the Columbia River and the age-composition of those returns is relatively high due to extensive sampling of Columbia River fisheries and recoveries of coded-wire tags (CWT). The forecasting methods for the Spring Creek Hatchery fall Chinook are identical to those reported above for Columbia River upriver bright fall Chinook.

In his workshop presentation, Steve Haeseker of the USFWS reported that natural variability in age composition makes forecasting of this stock difficult, but recent forecasts have been relatively precise and unbiased. For 1980-2015, mean percent error (MPE) was 8% and mean absolute percent error (MAPE) was 31% (Figure 8). Considerable under- or over-estimates (up to about +60%) of forecasted abundance have frequently occurred since the mid-1990s -- more frequently than with the upriver brights described above. Overestimates of abundance were commonly forecasted from 2006-2011, but 2013 was a substantial underestimate.

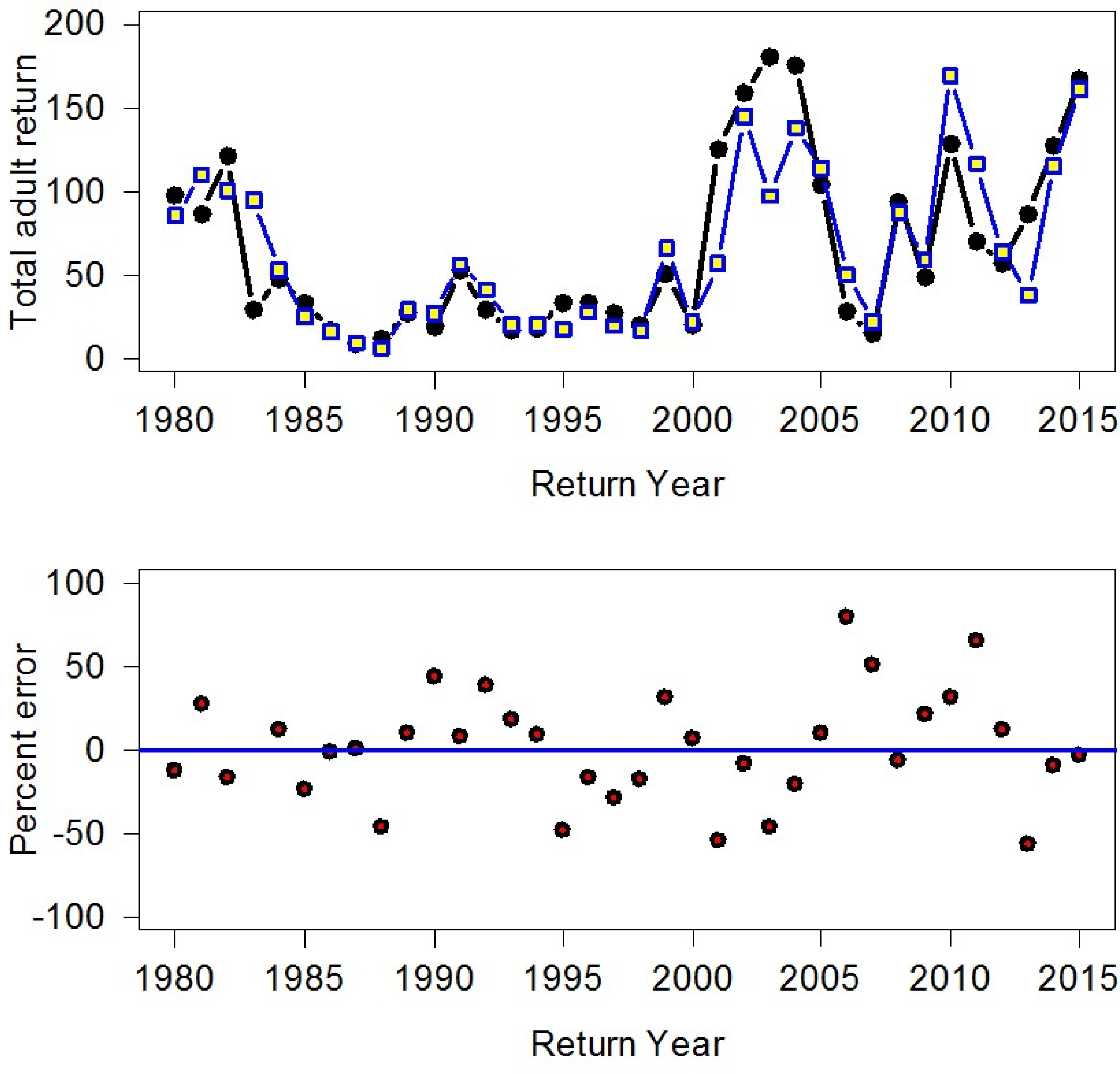


Figure 8. ***For return years 1995-2015, post-season estimates of abundance (thousands, black dots) and pre-season forecasts (yellow squares) (top panel), and percent error, ([forecast – actual]/actual) *100, between the pre-season forecasts and post-season abundance estimates for Columbia River mouth returns of Columbia River Spring Creek Hatchery Chinook Salmon (Steve Haeseker, workshop presentation). Positive errors mean the forecast was higher than the actual return.***

A major contributor to the frequent large forecasting errors for this stock over the last 20+ years is likely the large between-year and decadal-scale changes in age composition (Figure 9). Precision of these forecasts (based largely on sibling relations and cohort ratios) will necessarily be reduced when there are such large changes in proportions of age 4s and 5s between years, which create large prediction intervals.

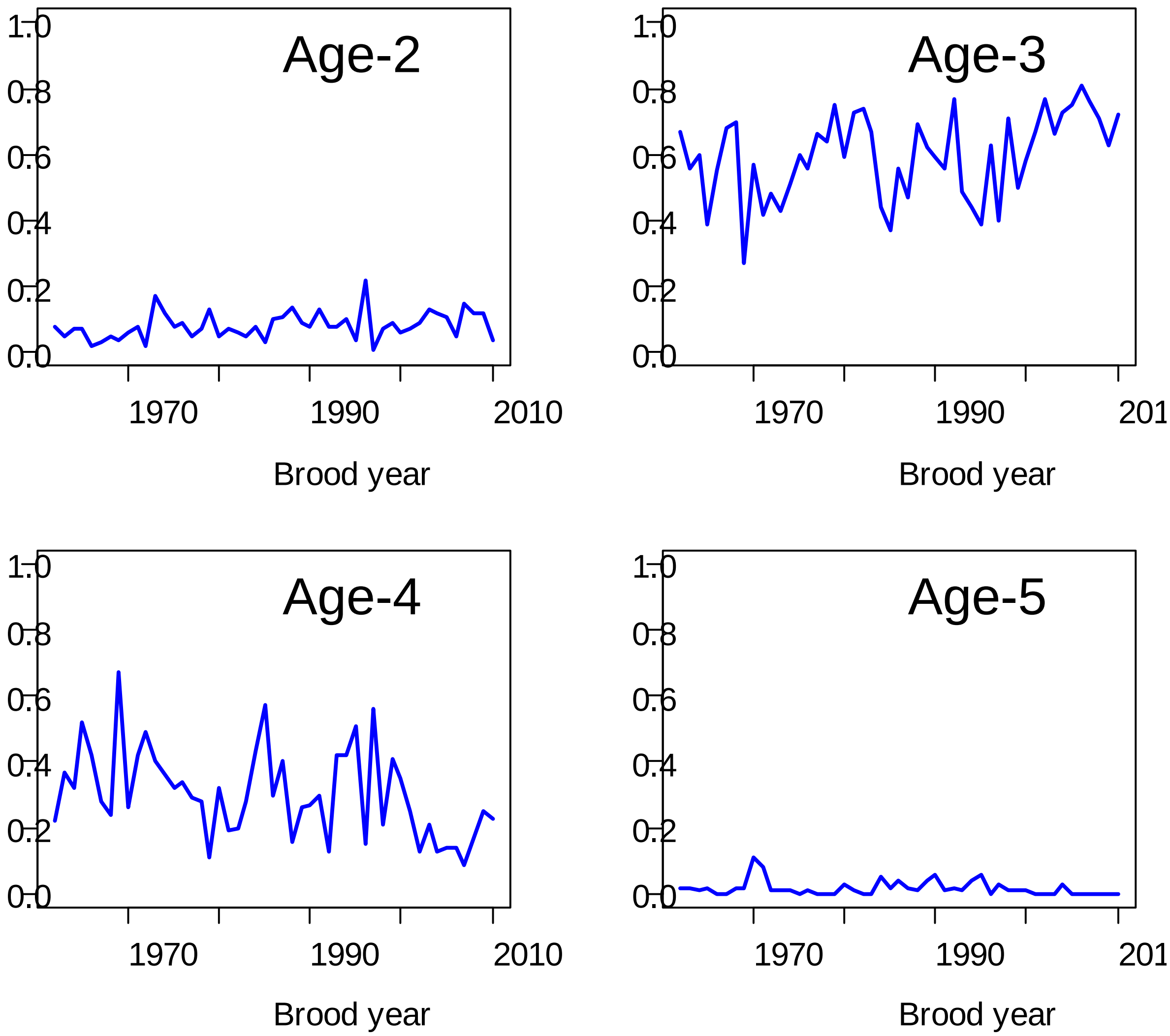


Figure 9. Proportions (Y axis) of different age classes of Columbia River Spring Creek Hatchery Chinook Salmon returning to the Columbia River for brood years 1962-2010 (Steve Haeseker, workshop presentation).

4.6 Columbia River Summer Run Forecast

Stuart Ellis (CRITFC) reported that the Upper Columbia River summer stock includes a mix of hatchery and wild fish produced in areas upstream of Priest Rapids Dam and the Yakima River. The recent 10-year average return of this stock to the Columbia River is approximately 71,000 (range 37,000 to 127,000). On average 13% return at age 3, 46% at age 4, 38% at age 5, and 3% at age 6 (Jeff Whisler, ODFW, personal communication).

CTC reports of the PSC Chinook model outputs estimate that substantial harvest of this stock occurs in AABMs and ISBMs from Southeast Alaska to the Columbia River (Figure 10).

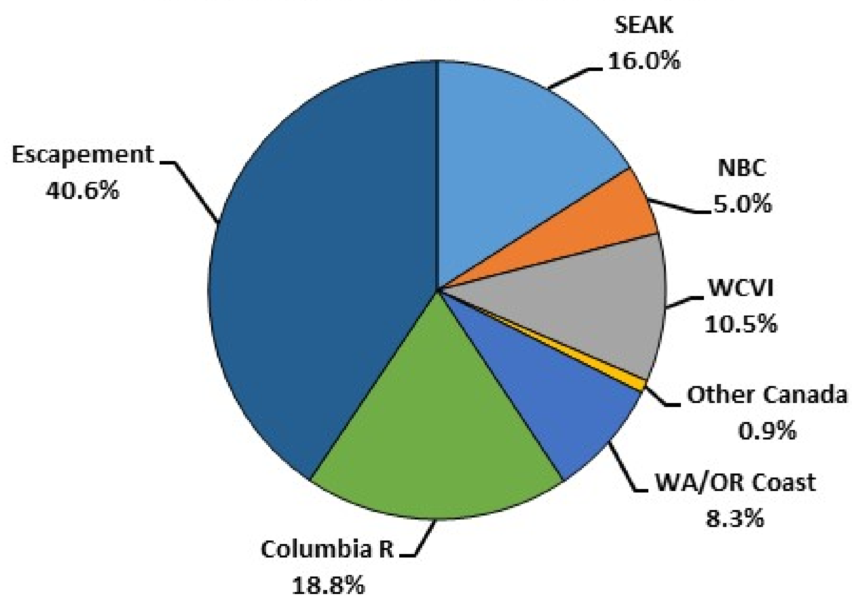


Figure 10. Distribution of Columbia River summer Chinook Salmon mortalities among fisheries and escapement, 1999-2013 (CTC 2015b).

Annual forecasts of Columbia River summer Chinook returns are produced by the U.S. v. Oregon Technical Advisory Committee (TAC). The quality of the data on total returns to the Columbia River and the age-composition of those returns is relatively high due to extensive sampling of Columbia River fisheries and recoveries of coded-wire tags (CWT). The primary forecasting methods are arithmetic-scale sibling regressions and average cohort ratios (e.g., the average ratio of age-3 to age-4 returns from the same brood year). The TAC typically provides point-estimate forecasts based on the age-specific best-performing year ranges of input data identified by consensus.

In his workshop presentation, Stuart Ellis of CRITFC reported that natural variability in age composition makes forecasting of this stock difficult, but recent forecasts have been relatively precise and unbiased. For 2005-2015, mean percent error (MPE) was 5% and mean absolute percent error (MAPE) was 24% (Figure 11). The run was forecast too high from 2009-2012 but too low in 2014 and especially 2015.

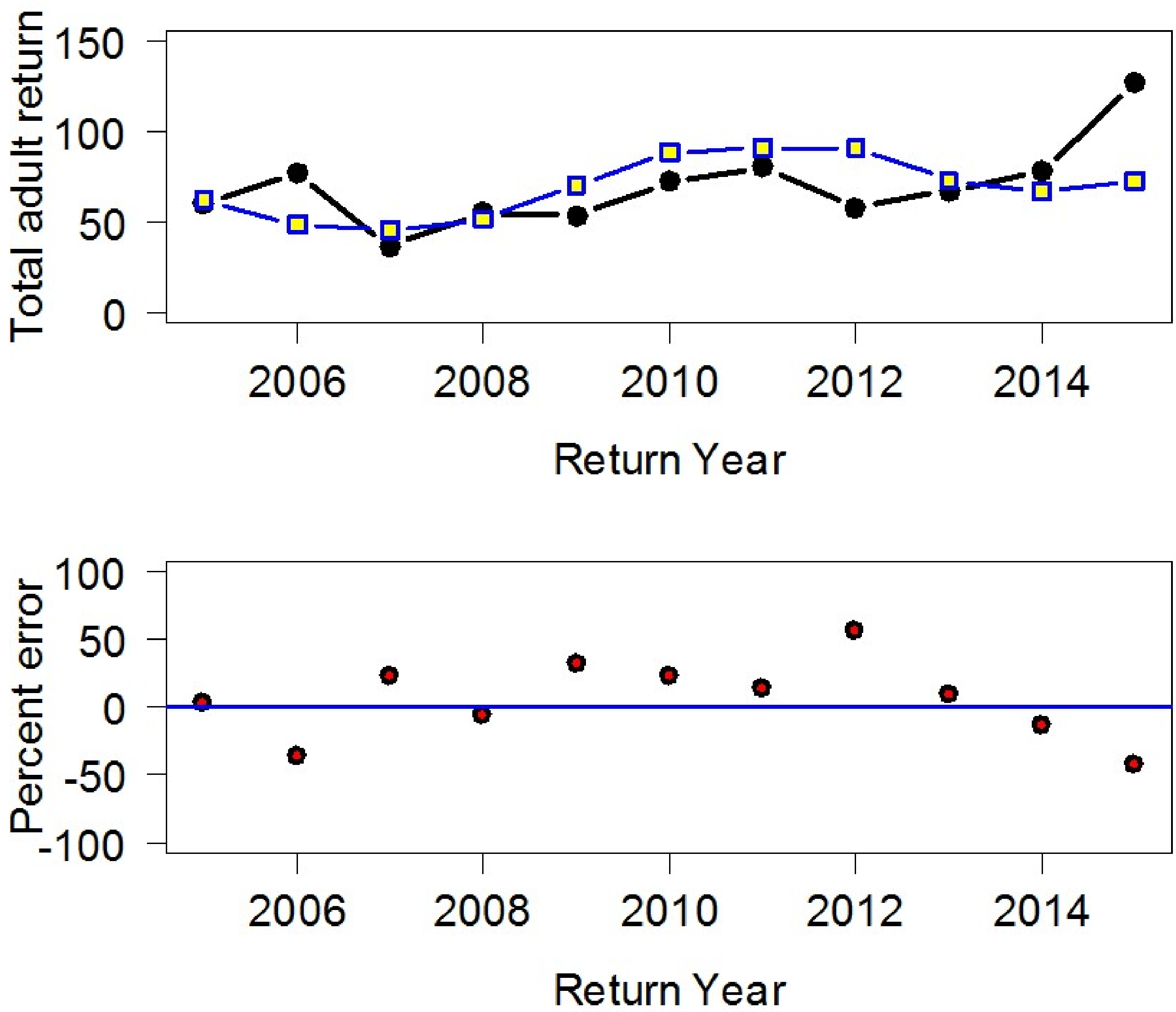


***Figure 11.** For return years 2005-2015, post-season estimates of abundance (thousands, black dots) and pre-season forecasts (yellow squares) (top panel), and percent error, ([forecast – actual]/actual) *100, between the pre-season forecasts and post-season abundance estimates for Columbia River mouth returns of Upper Columbia summer Chinook Salmon (Stuart Ellis, workshop presentation). Positive errors mean the forecast was higher than the actual return.*

4.7 Northern Oregon Coast Fall Forecast

Ethan Clemons (ODFW) reported that the Northern Oregon Coast (NOC) "stock" is an aggregate of populations returning to small rivers including the Siuslaw, Alsea, Yaquina, Siletz, Salmon, Nestucca, Tillamook and Nehalem. The total aggregate return has varied from about 40,000 (1970s, 2008) to over 170,000 (1988, early 2000s). Age at maturity is typically 3 to 6 years with a small component of 2-year olds.

CTC reports of the PSC Chinook model outputs estimate that substantial harvest of this stock occurs in AABM fisheries of Southeast Alaska and Northern British Columbia as well as terminal fisheries of the Oregon coast (Figure 12).

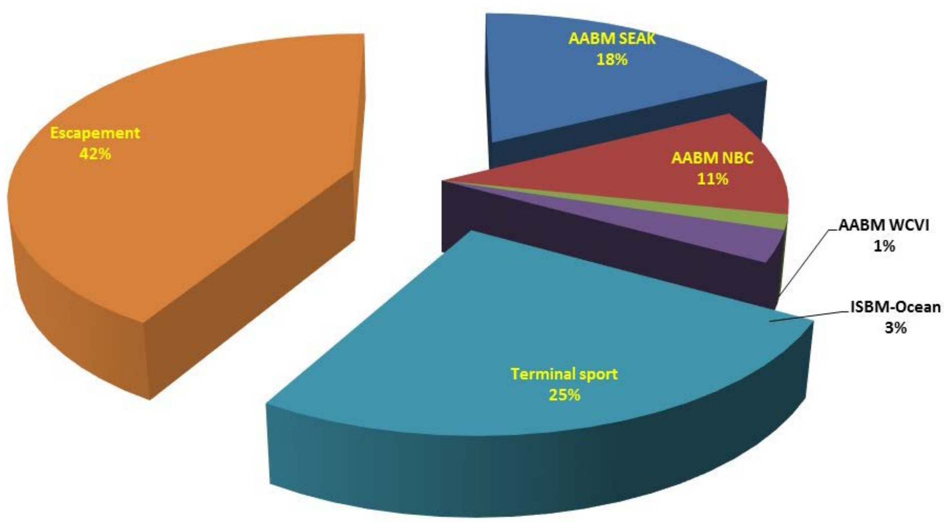


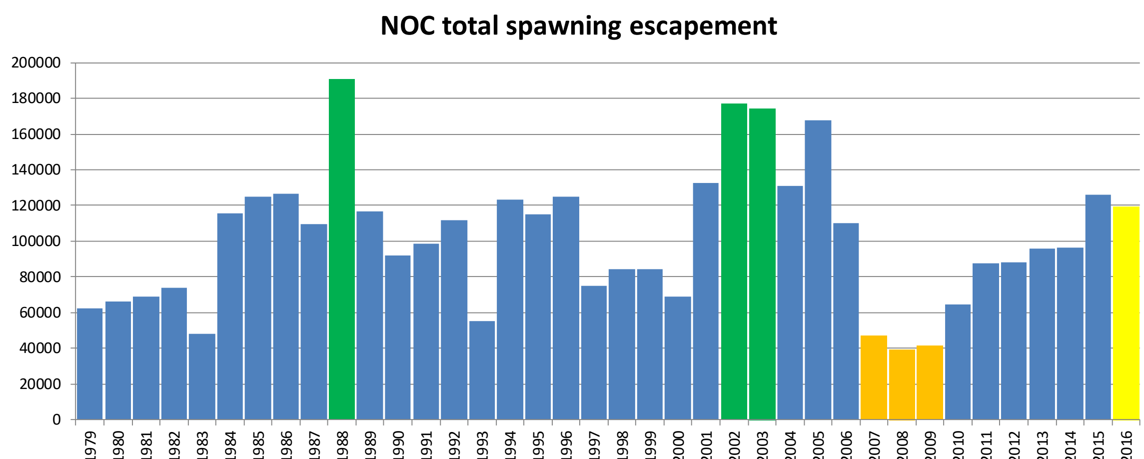
Figure 12. Distribution of North Oregon Coast Chinook Salmon mortalities among fisheries and escapement, 1999-2014 (CTC 2015b, Ethan Clemons ODFW unpublished data).

The 2008 PST renegotiation highlighted data limitations for stock forecasting for the NOC aggregate. Forecasts at that time were based on a 3-year average of escapement. Spawner index surveys were being conducted, but age sampling was limited and scales were not read in time for use in forecasting. A precipitous decline in escapement from 2007-2010 drew additional management attention to NOC Chinook by ODFW, which led to a change in agency priorities and rapid turnaround of scale aging data starting in 2008. This allowed forecasting of annual returns based on sibling regressions. Forecasting methods have been refined since 2008 based on each year's forecast performance. Several different sibling-regression relationships have been considered, and no single method has been consistently applied to all stocks in all years. In 2016, ForecastR modules were used. They allowed development of ARIMA models for some stocks in time for use for forecasting input to the PSC model.

Forecasts are for escapement only, not pre-ocean-fishing abundance or terminal returns. Current models assume that all fisheries are going to have the same proportional impact as they have had during the years that were used to generate the sibling relationships and time series models. It is currently not practical to forecast the varying impact of AABM or terminal fisheries in a usable time frame.

Annual escapement is estimated based on spawning ground surveys. Forecasts are generated for each of seven populations and then aggregated into the NOC stock. Expansions are made for unsurveyed areas assuming a static relationship between surveyed streams and unsurveyed streams/basins (expansion by 17%). Maturation rates and year-to-year survival rates are assumed to be static. Age-specific sampling is assumed to be unbiased (or corrected for known biases).

In his workshop presentation, Ethan Clemons of ODFW) reported that recent forecasts have been relatively precise and unbiased (Figure 13). However, funding reductions have substantially reduced spawning ground survey effort in recent years, so the current quality of stock assessment will not be sustained.



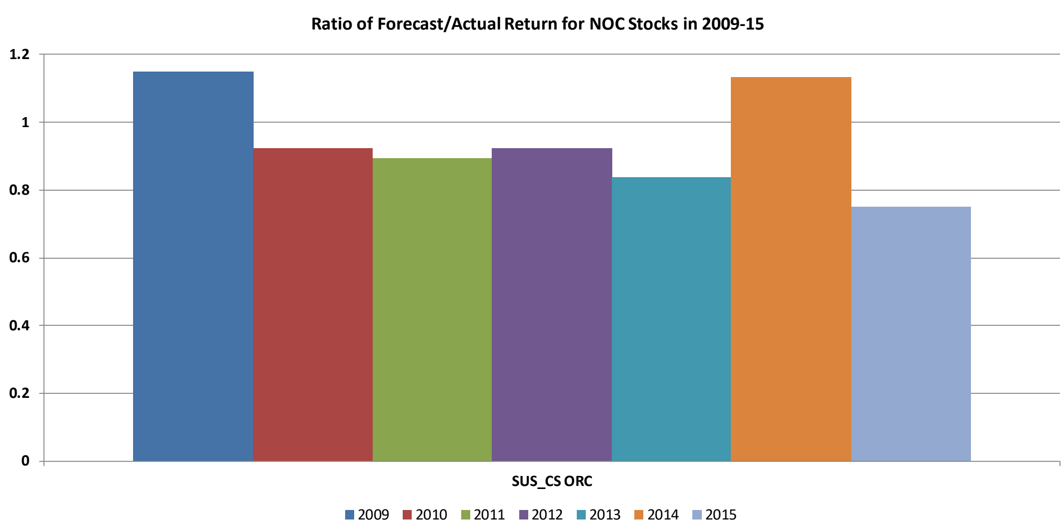


Figure 13. Spawning escapement and forecast/actual escapement data for Northern Oregon Coast Chinook Salmon (Ethan Clemons, ODFW, workshop presentation).

5 Conceptual Framework for the Panel's Work

The Panel conducted its review through the “lens” of a general conceptual framework of sources of uncertainty in forecasting methods. Uncertainty is manifested in both accuracy and precision of forecasts. Here "accuracy" is a measure of how close an estimated value is to the "true" value. If repeated estimates over time are consistently too low or too high, they are statistically biased, or inaccurate. "Precision" describes how similar multiple estimates are to each other, regardless of their bias.

If everything were known perfectly, then there would be no forecasting errors, but of course, that is impossible. Scientists have an incomplete understanding of the dynamics of salmonid population dynamics, ecosystems, and fishing dynamics. The resulting uncertainties are reflected in assumptions and hypotheses embedded in the statistical models of regional forecasting agencies and in the PSC's Chinook model. These uncertainties can be grouped into four categories, (1) unclear management objectives, (2) structural uncertainty, (3) uncertainty in parameters, and (4) outcome uncertainty. We define these categories here and give examples of each in sections 6, 7, and 8, along with recommendations for how to deal with them.

5.1 Unclear Management Objectives

Quantitative fisheries analysts know that in order for their analyses to be directly useful to fisheries managers, the calculated indicators of fish stocks and fisheries should fit into clearly articulated management objectives. To choose an extreme hypothetical example, if managers were most concerned about the chance of low salmon abundance occurring during the next 5 years, then it would be inappropriate for modelers to merely show the long-term average abundance expected over that period. Instead, given that management objective, analysts should calculate indicators of frequency and magnitude of abundances below the managers' undesirable level.

This point seems obvious, but it is surprising how often the uncertainty caused by the lack of clearly stated operational management objectives leads to inappropriate scientific advice and/or confusion on the part of scientists and managers. Participants at the Portland workshop expressed this need for clear management objectives, both for stock-specific forecasts made by agency scientists and for forecasts of abundance in AABMs made by PSC modelers. Only with such clarity will forecasting models produce output that directly meets the needs of decision makers.

5.2 Structural Uncertainty

Structural uncertainty refers to the lack of certainty about which equations in a model are correct (i.e., reflect reality). If only one form of an equation in a forecasting model is used to represent a given process (for example, a linear instead of nonlinear sibling relation), then implicitly the analyst is saying that the probability is 1.0 that the underlying natural process is linear and that the probability is zero that any other shape of function exists. Similarly, the assumption in the PSC model that Chinook stocks have the same oceanic spatial distribution now that they had during the 1979-1982 base period implies that there is zero probability that the distribution has changed. Unfortunately, we may be wrong about these assumptions because we have incomplete knowledge of the real world. If those assumptions are indeed wrong, then the single point estimates of forecasts of Chinook abundance are also likely to be wrong. Such point estimates would therefore not reflect the structural uncertainty in forecasts.

5.3 Parametric Uncertainty

Uncertainty in model parameters refers to the lack of certainty about quantitative values such as natural mortality rate between ages, maturation rates, exploitation rates, as well as parameters of spawner-recruit models, sibling relationships, or other equations. Such parameters are estimated through various means, but those estimates are likely to differ from the true underlying values because of natural variability in processes that are not fully described by the equation (e.g., spawner-recruit model) and/or observation error in stock composition, abundance of spawners, and catches. Such errors in parameter values in agency and PSC models will lead to errors in forecasts of abundance. The magnitude of forecasting errors will depend on which parameters are wrong in which equations.

5.4 Outcome Uncertainty

Outcome uncertainty is a broad term referring to the deviation between some management target and the actual realized outcome (Holt and Peterman 2006). For instance, it refers to the difference between a maximum allowable catch (e.g., 150,000 fish) and the actual catch (e.g., 200,000), or between a target harvest rate of 40% and the actual outcome of 30%, or between an escapement goal of 50,000 and actual escapement of 40,000. Such deviations can arise from at least five sources: (1) the vulnerability of fish to fishing gear (catchability) differed from the expected level because of unexpected changes such as body size, depth, or horizontal location of the fish, (2) non-compliance by fishing vessels with regulations (sometimes referred to as imperfect control, implementation uncertainty, or implementation error), (3) errors in forecasts of abundance, (4) errors in post-season estimates of abundance or catch, and (5) management regulations that were not the correct ones to meet the objectives, even without the problems of sources (1) and (2). Outcome uncertainty is relevant to agency as well as PSC model forecasts of Chinook abundance because both make pre-season assumptions about exploitation rates in AABM fisheries that won't occur until after the forecasts are made.

5.5 Implications and Perspective

These four sources of uncertainty provided a useful way for the Panel to organize its review and to develop recommendations. Explicitly defining the types of forecast uncertainties will pave the way for both identifying measures to reduce them and accurately reflecting them in forecast estimates. The Panel recognizes that regional agency forecasters, as well as CTC modelers (the Analytical Working Group, or AWG), are very experienced quantitative scientists who are already very well aware of these types of uncertainty. Nevertheless, consideration of these uncertainties are not always clearly articulated in their annual reports. For instance, regional agencies produce annual stock-specific point-estimate forecasts that omit uncertainties, and these are sent to the CTC modelers for input to the PSC model. Similarly, the deterministic PSC model produces point estimates of forecasts for use in Table 1 of Chapter 3 of the Treaty.

The Panel recognizes that forecasting abundances of North Pacific salmon populations is difficult, even with the best of data sets and methods, so we begin by placing the forecasts of Chinook Salmon into a broader context. To our knowledge, no synthesis has been conducted that quantitatively compares management agencies' historical pre-season forecasts with actual returns across all major Pacific salmon species, stocks, and areas. However, an almost equally informative analysis was reported by Haeseker et al. (2008), who compared actual returns across decades with forecasts that would have been made in each historical year if those forecasts had been based on the best of 11 statistical forecasting models for each chum salmon stock and each sockeye salmon stock. We see no inherent reason why Chinook Salmon forecasting should be any better or worse than that for sockeye salmon, which shares with Chinook the tendency to mature at 3 or more ages. Chum salmon are also relevant, but most often chum stocks only have two ages at maturity.

The best or top-ranked model based on MPE (long-term statistical bias) varied considerably among chum and sockeye salmon stocks, ranging from sibling models to naïve averages of recent returns, as it did for MAPE (precision of forecasts) (Haeseker et al. 2008). *The top-ranked model for each stock* produced an average MPE of 19% across 40 chum salmon stocks and 27% across 37 sockeye salmon stocks (Haeseker et al. 2008 and spreadsheet). An MPE of zero is the most desirable value, representing no statistical bias.

The frequency distributions of these stock-specific MPE values for chum salmon and sockeye salmon show that the MPE values for all three Columbia River stocks plus the Northern Oregon Coastal stock are at the low end of the range of MPE values observed in chum and sockeye salmon (Figure 14). Given this perspective, the forecasts for these four stocks are doing well, but there is still some possibility of improvement. In contrast, with an MPE of -26.9%, the WCVI forecasts fell below the range observed for either chum or sockeye, and below all four other Chinook stocks examined here. Clearly some substantial improvement is needed fro WCVI.

On a related point, we note that the CTC's expectation of MPE of $\pm$7.5% for forecasts of terminal runs or escapements (CTC 2016b) may be too stringent. The Panel learned that the CTC also sets minimum data-quality standards for escapements and catches, which in principle is fine, but over-emphasis on those standards may be misplaced given the large number of other sources of uncertainties described in this report that are not taken into account in either the agencies' forecasting models or the PSC model.

For the measure of forecasting precision, MAPE, *the top-ranked model for each stock* produced an average MAPE of 58% across 40 chum salmon stocks and 66% across 37 sockeye salmon stocks (Table 1) (Haeseker et al. 2008 and spreadsheet). The frequency distributions of these stock-specific MAPE values for chum salmon and sockeye salmon show that the MAPE values for all three Columbia River stocks plus the Northern Oregon Coastal stock are at the low end of the ranges observed in those other two species (Figure 15), which again reflects good performance for those models, but with some room for further improvement. The MAPE of 42% for the WCVI forecast is higher than the other four Chinook stocks examined here but is still well within the range of MAPE values for chum and sockeye salmon.

Table 1. Mean percent error (MPE, bias) and mean absolute percent error (MAPE, precision) for forecasts produced by regional agencies for the five Chinook Salmon stocks considered in this review. Also shown are average MPE and average MAPE values for the best (closest to zero) stock-specific model for each of 40 chum salmon stocks and 37 sockeye salmon stocks (Haeseker et al. 2008). Data cover return years 1999-2015[a] for the five Chinook Salmon stocks and from as far back as 1974 through 1999 return years for the chum salmon and sockeye salmon stock. Mean percent error was calculated by [(forecast-actual post-season)/actual post-season] times 100.

Stock	Mean Percent Error (MPE)	Mean Absolute Percent Error (MAPE)
West Coast Vancouver Island Chinook	-26.9%	42.1%
Columbia River Summer Chinook	9.7%	22.2%
Columbia River Spring Creek Fall Chinook	7.6%	31.3%
Columbia River Upriver Bright Chinook	0.9%	25.1%
Northern Oregon Coastal Fall Chinook	8.4%	31.3%
Best stock-specific model for each of 40 chum salmon stocks (Haeseker et al. 2008)	19% (median 12%)	58% (median 52%)
Best stock-specific model for each of 37 sockeye salmon stocks (Haeseker et al. 2008)	27% (median 15%)	66% (median 57%)

[a] *Data for 1999-2013 from CTC (2015b, Appendix J1) and preliminary data for 2014 and 2015 obtained from John Carlile (ADF&G, personal communication, 9 Sept. 2016)*

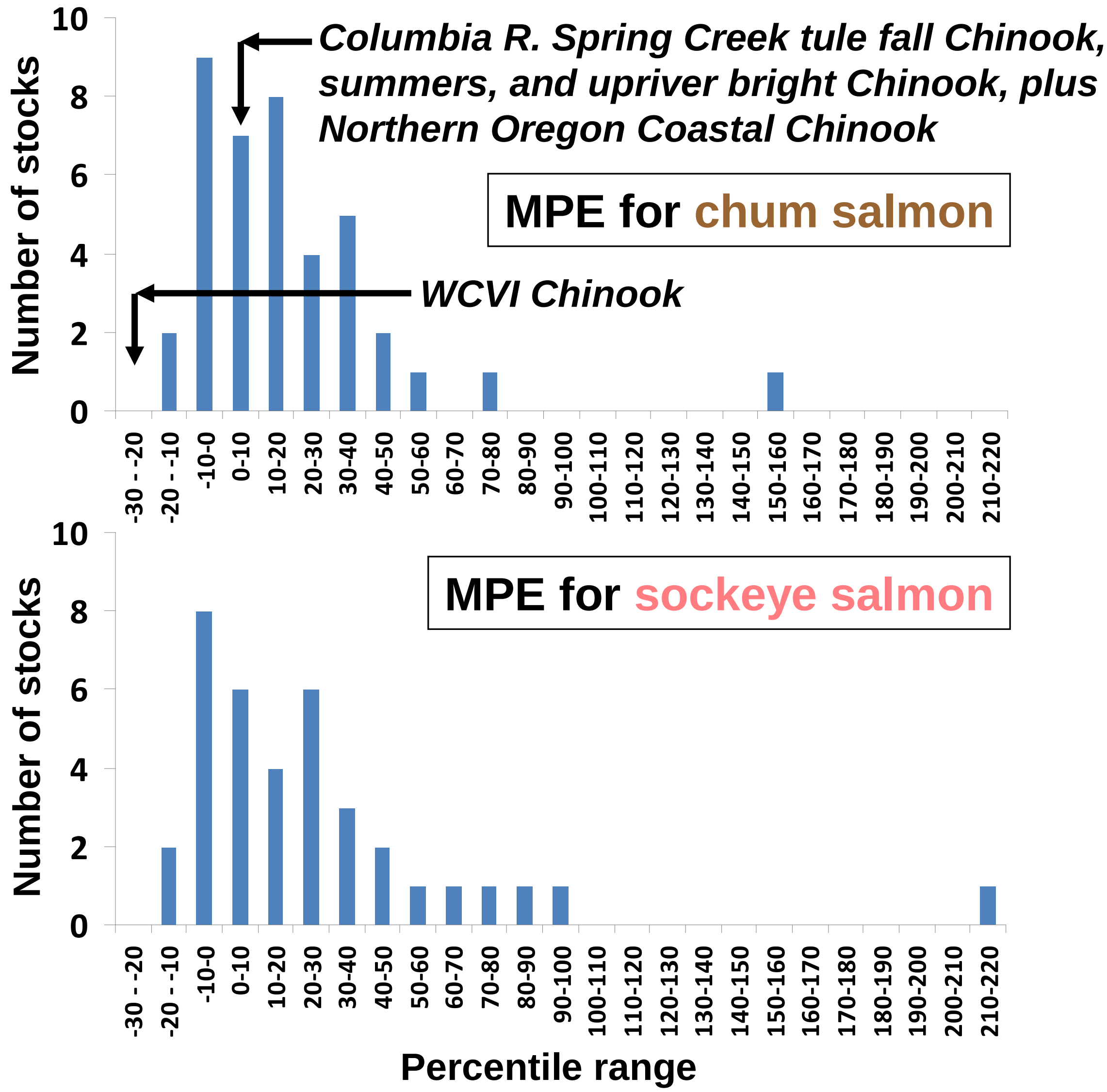


Figure 14. Frequency distributions of of a measure of bias of forecasts, mean percentage error, MPE, [(forecast-actual post-season)/actual post-season] times 100, for the top-ranked model for each stock across 40 chum salmon stocks (top panel) and 37 sockeye salmon stocks (bottom) (Haeseker et al. 2008). The percentile range for the five Chinook stocks examined in this review are shown by arrows in the top panel. Forecasts that are biased high (overestimates) have positive MPE values.

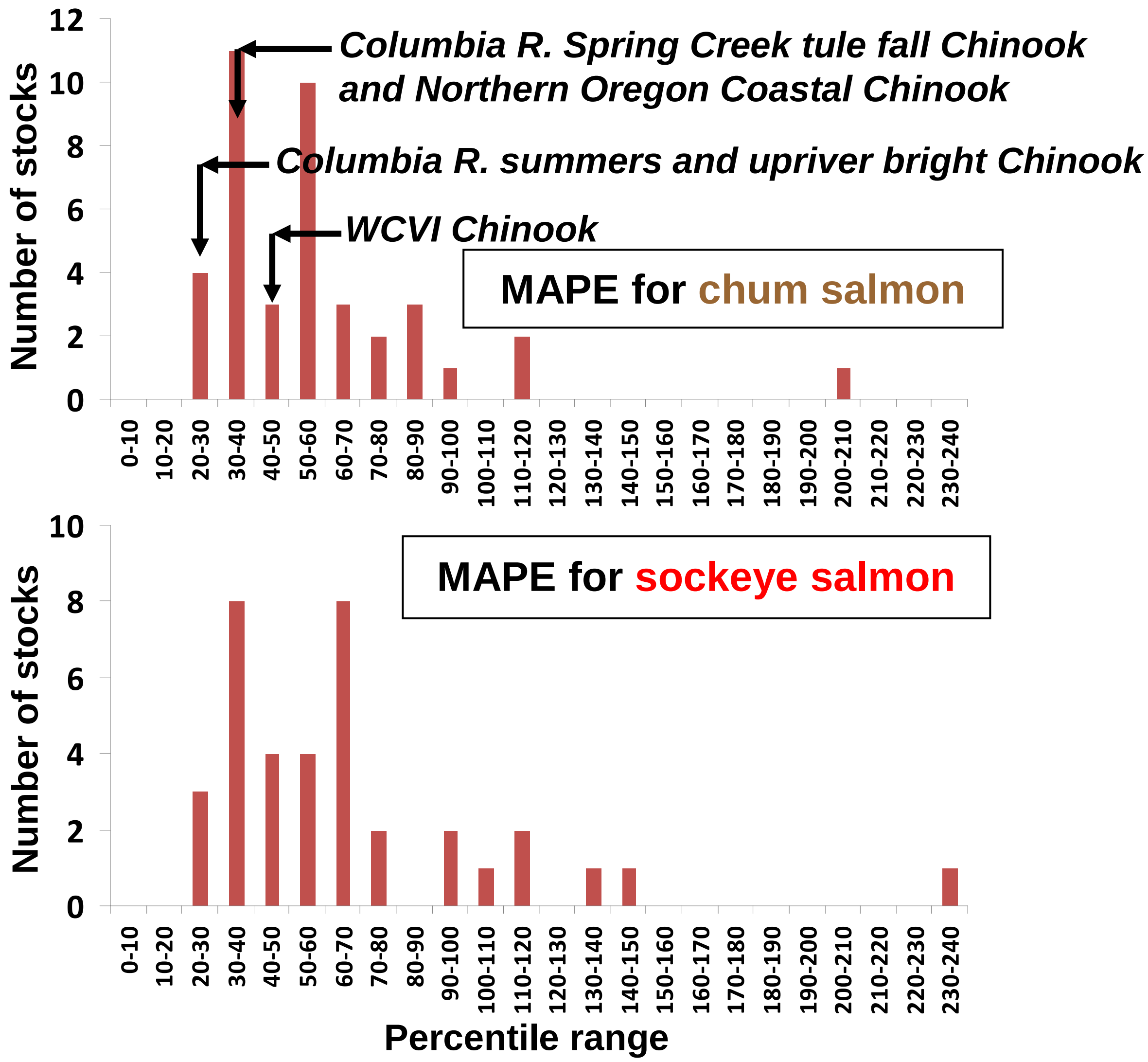


Figure 15. Frequency distributions of a measure of precision of forecasts, mean absolute percentage error, MAPE, for the top-ranked model for each stock across 40 chum salmon stocks (top panel) and 37 sockeye salmon stocks (bottom) (Haeseker et al. 2008). The percentile range for the five Chinook stocks examined in this review are shown by arrows in the top panel.

6 General Issues and Conclusions

The Panel identified several thematic issues that apply to forecasts of abundance of specific Chinook Salmon stocks as well as the use of the PSC model. These general issues and conclusions are important enough to the Treaty process that they warrant emphasis at the start of this report. More detailed recommendations consistent with these conclusions are found in section 7 on agencies' stock-specific forecasts and section 8 on the PSC model's forecasts of abundance in AABM areas.

6.1 Documentation of Agency Forecasting Methods and Results

Issue I. Current documentation of agency forecasts of abundance that are sent annually to the CTC does not provide sufficient information for PSC modelers to identify the long-term accuracy and precision of those forecasts, let alone uncertainty about the current year's forecast.

As noted in Issue 1 above, biologists for the five stocks that are included in this Panel review, produce annual forecasts of abundance for their own within-region management advice, as well as for input to the PSC Chinook model for estimating pre-fishing ocean abundance indices in AABMs. Agency forecasts are produced by a variety of methods, depending on the stocks and years, as is described in more detail in sections 4 and 7.

While the PSC modelers cannot currently incorporate information about uncertainties in the agencies' forecasting methods directly into their model, they expressed the desire to consider the accuracy and precision of forecasts provided by the regions when reviewing data inputs to the PSC model. Pertinent information should include not just the point estimates of forecasts, but also details of alternative forecasting models that were considered, the basis for selecting the final forecasting model, its critical assumptions and uncertainties, its long-term accuracy and precision documented via a retrospective analysis (defined in section 6.4), and a measure of uncertainty about the current year's forecast.

The Portland workshop also identified discrepancies between numerical values of certain agency forecasts reported by the CTC and forecasts that were originally submitted to the CTC by agency staff. It is unclear whether this was due to incomplete documentation of updated agency forecasts or other issues. Regardless, such errors can be avoided by agency representatives assuming responsibility for both documenting their submissions and reviewing CTC reports to ensure that their information was applied and reported correctly.

Conclusion A. More comprehensive documentation is needed by the CTC from regional agency forecasters regarding the agencies' methods, critical assumptions and uncertainties, and accuracy and precision of past stock-specific forecasts. Agencies should also state the uncertainty in each stocks' annual

forecasted abundance. More frequent in-depth communication between PSC modelers and agency staff is also required.

6.2 Requirements for Stock Forecasts as Inputs to the PSC Chinook Model

Issue II. Efforts by agencies to provide forecasts as inputs to the PSC model are hampered by an incomplete understanding of (1) the PSC model's information requirements, (2) how those forecasts are used in that model, and (3) how those uses differ from those of fishery managers within regions.

Annual forecasts of stock-specific abundance are typically generated by staff in fishery agencies for terminal runs or escapements to regional management units (e.g., Columbia River, Northern Oregon Coast, or West Coast Vancouver Island). One purpose of agency forecasts is to plan and configure terminal area fisheries to meet established escapement goals, catch or exploitation-rate limits, and allocation objectives in those local areas. In addition, terminal run size/escapement forecasts are also sent through the CTC to PSC Chinook modelers who are tasked with taking those terminal run size/escapement forecasts into account and producing pre-ocean-fishery abundance indices (AIs) for use in establishing exploitation rates and corresponding catch limits in the three AABM fisheries. In some cases, forecasts for use within regions for management of domestic fisheries differ from forecasts sent to the CTC for the PSC model, and that difference may be quite appropriate because they are intended to be used for different purposes.

However, it was apparent at the Portland workshop that communication between PSC modelers and regional agency forecasters is often incomplete. PSC modelers have specific requirements for annual forecasts, but they are not formally documented and transmitted to agency forecasters. As well, many agency scientists have a limited understanding of how their forecasts are being used in the PSC model.

Potential points of confusion range from relatively simple questions (whether forecasts should include or exclude jacks) to more complex issues (assumptions of appropriate marine exploitation rates for forecasting terminal returns). The WCVI forecast application in AABM fisheries was a particular concern. For instance, it was unclear for the WCVI forecaster and the CTC modeling group, as well as in the PSC model's documentation, how Fishery Policy (FP) adjustments (i.e., scalars to the exploitation rates) have accounted for the change in magnitude and spatial distribution of fishing effort that has occurred since the 1979-1982 base period. As a result, until recently WCVI forecasts have apparently been generated using base-period exploitation rates, which may have been substantially greater than current fishing rates. If this reduction in recent exploitation rates is indeed correct, then it is no surprise that those forecasts of pre-season abundance that were based on base-period exploitation rates have chronically underestimated WCVI Chinook abundance. In recognition of this problem, starting in 2014, separate WCVI forecasts were also generated with exploitation rate assumptions that reflect recent fisheries, and both types of WCVI forecasts were sent to the CTC for input to the PSC model.

Annual written requests from the CTC to agency forecasters should identify whether agency forecasts should include jacks or not, what units/currency to use for terminal runs and escapement, and whether to use exploitation rates from recent years or the base period.

Conclusion B. More explicit direction from the Chinook Technical Committee is needed by agency-based stock forecasters regarding the annually requested forecasts.

6.3 Limitations of Existing Stock Assessment Data

Issue III. Accuracy and precision of stock forecasts are limited by the available stock assessment data; this is more of a problem for some Chinook stocks than others.

At the workshop, the Panel heard several concerns about the quality of data for escapements, age structure, and harvest, as well as how they were being used. These concerns are not new but are important to highlight in view of the expectations of the Treaty's abundance-based management framework and the corresponding use of abundance indices produced by the PSC Chinook model. Accurate forecasts of abundance are essential for effectively implementing abundance-based management. A high potential for measurement error in data fed into the stock-specific and PSC models substantially reduces the ability to make those forecasts accurate. Another concern is that demands for increasing model specificity (such as the ambitious current plan to expand the PSC's model from 25 to 48 fisheries and from 30 to 40 stocks) may easily surpass the quality of the available data.

Finite resources for stock assessment are always a challenge and contribute to substantial uncertainties associated with forecasts for several stocks. For example, the quality of stock assessments of Northern Oregon Coast Chinook stocks appear to have been substantially upgraded in recent years from historical levels. However, current funding for those assessments is expected to decrease, which may substantially reduce the accuracy and precision of future estimates. West Coast Vancouver Island Chinook assessments are similarly hampered by the need to make significant inferences from very limited data. As the Panel understands it, the latter assessments currently rely on Robertson Creek Hatchery data, which are then expanded to represent other WCVI hatcheries and wild-stock production. Uncertainties in this WCVI process are potentially very significant and may lead to substantial over- or underestimation of stock status, including wild and hatchery abundance.

In addition to specific stock concerns, substantial uncertainty is introduced by more systemic limitations of existing information that are already widely acknowledged by the CTC and agency forecasters. Chief among these is the assumption that exploitation rate and marine survival rate of wild fish are identical to those derived from coded-wire-tagged hatchery Chinook indicator stocks. The potential for consistent underestimation of stock contributions to catches based on a low incidence of CWT observations in some fisheries must also be acknowledged.

Better data are needed for several stocks to support high expectations for forecast accuracy and precision. Conversely, expectations might need to be scaled down to reflect existing uncertainties.

Conclusion C. Substantial improvements in basic assessments of some Chinook stocks are needed to support current PSC model and management applications, otherwise expectations need to be rescaled/reduced to recognize existing data limitations. Further expansion of the PSC model's number of stocks and fishing areas may need to be postponed until the quality of relevant data is deemed suitable.

6.4 Definitions and Best Practices for Agency Stock Assessment and Forecasting

Issue IV There is substantial differences among regional agencies in how stock forecasts are produced and described.

Stock forecasting methods are tailored to the specifics of the information, past practices, and experience of forecasters in each region. Forecasts rely heavily on sibling relationships and average ratios of successive ages in successive years, and all agencies have explored various years of data sets for estimating parameters. It would be counterproductive to try to impose a single standard of forecasting practices across regions, but a set of standard definitions and best practices could be a helpful reference to improve the statistical foundation of methods for stock forecasting by agencies. Many decades of experience by fisheries scientists has led to a set of common practices in fish stock assessment that have proven to be effective.

A few examples of definition and practices identified by the Panel include:

Resolution of forecasts: Annual age- and sex-specific estimates of total escapement should be available. Point estimates should be accompanied by estimates of uncertainty.

Expansion factors: The source of expansion factors from index values to larger aggregates should be documented and some measure of interannual variability of those factors should be quantified.

Measures of forecasting errors: To facilitate comparisons of forecasting errors across stocks and models, the CTC should agree upon a minimum set of standard measures of those errors that should be produced by all agencies as well as the CTC. This simple step will eliminate the current inefficiency and confusion caused by the use of several different measures of forecasting errors in different documents, which preclude direct comparisons. Such diverse measures for stock-specific forecasts include (1) [(forecast - actual)/actual], (2) [(actual - forecast)/actual], (3) forecast/actual, (4) actual/forecast, and (5) some of those multiplied by

100 and others not, etc. For abundances produced by the PSC model, the analogous formulas would use pre-season AIs instead of "forecast" and post-season AIs for "actual". We suggest the format of option 1 above as a default so that overestimates by a forecasting model are displayed as positive values and underestimates as negative values.

Retrospective analysis: The strongest test for evaluating the performance of alternative forecasting models (short of using Monte Carlo simulation models) is to conduct a retrospective analysis. This procedure forecasts abundance for a given historical year based only on data that would have been available up to that year, and then iteratively repeats this process after adding that year's actual returns, and works through the time series of data (see section 7). Limitations of retrospective analyses of alternative models also need to be recognized, though. Model rankings may be affected by (1) which particular years were used to initially fit the model prior to the first forecast, (2) the length of time series used to calculate performance, and (3) the nature of historical variability (whether it will likely encompass future situations).

Align model-ranking criteria with management objectives: Most agencies use more than one ranking criterion or "performance measure" each year for choosing the best forecasting model, and these criteria can differ among years and stocks. However, these ranking criteria implicitly reflect different management priorities, so due diligence needs to be paid by agency forecasters to use model-ranking criteria that provide the most appropriate information (i.e., that is consistent with stated management objectives for both specific stocks and for AABMs). We provide some hypothetical examples below.

Mean raw error (MRE) (the average of positive and negative forecasting errors over many years) and its scaled counterpart, mean percent error (MPE), measure the long-term bias in forecasts (i.e., how much on average a given model tends to over- or underestimate abundance across the entire period). Thus, use of MRE or MPE for choosing the best forecasting model would be appropriate for a management objective that is only concerned with whether there is a long-term tendency for a consistent bias in one direction or the other, i.e., consistently over-estimating the run or under-estimating it, and without any concern about the year-to-year variability in forecasting errors. To reflect a management objective focused on the latter variation, though, mean absolute error (MAE) and mean absolute percent error (MAPE) are appropriate because they measure the precision, or average magnitude of annual forecasting errors, regardless of whether they are over or under actual abundance. This ranking criterion fits with a management objective that puts top priority on coming as close to the actual post-season estimate of abundance as possible, regardless of the sign of the error. That objective implicitly places equal weight on over-forecasting by some amount (e.g., 100,000 fish) and under-forecasting by that same amount.

Other model-ranking criteria reflect management objectives such as minimizing both bias and precision (root-mean square error, RMSE) or more heavily weighting overestimates than underestimates (or vice versa). Another criterion, r^2, reflects the proportion of variation in year-to-year post-season abundances that is accounted for by a given forecasting model. Finally, when the new statistical forecasting package, ForecastR, is released, other formal model-ranking criteria will become readily available to Chinook forecasters, such as the Akaike Information Criterion for small samples (AIC_c).

Caution is advised, though. All of the model-ranking criteria mentioned above, including AIC_c, implicitly assume that forecasting errors of a given magnitude in one direction are as undesirable as the same magnitude of forecasting error in the opposite direction, but this may not be appropriate for some management objectives (p. 101 in Walters and Martell, 2004). For instance, if managers place different weightings on errors in different directions, then other model-ranking criteria will need to be developed instead.

In short, agency forecasters can provide forecasts from models that meet a variety of performance criteria, but in order to choose those criteria and produce numerical results that will be most useful to managers, management objectives need to be clearly stated in quantitative, measurable forms. Of course, such discussions between managers and scientists should be seen as a way to ensure that statistical analyses efficiently address management concerns. Such discussions are not intended to have scientists influence the choice of value-laden management objectives, nor to have managers influence assumptions or outcomes of scientific analyses.

Sibling regressions: $\log_e$-$\log_e$ model equations are more likely to meet assumptions of regression in sibling age-class models (Peterman 1982), but it is important to apply the usual log-normal bias correction when back-calculating abundance on the arithmetic scale (Haeseker et al. 2005). The ForecastR package described below contains a $\log_e$-$\log_e$ sibling model, and we were told that it includes the log-normal bias correction mentioned in the previous sentence.

Kalman filter estimation of sibling regressions: Sibling age-class relationships fit by standard regression assume constant age-specific maturation rates and survival rate between ages. However, large scatters of data points around some sibling relationships, as well as time trends in mean-age-at maturity that have been observed in Chinook Salmon (CTC 2016a) and sockeye salmon (Pyper et al. 1999), suggesting that better forecasts might be possible by fitting sibling models using a Kalman filter estimation procedure (Holt and Peterman 2004). When a sibling model is set up to be estimated via a Kalman filter, the procedure estimates temporal changes in parameters of sibling age-class relationships and takes into account observation error as well as natural variability. This procedure has already proven effective for sockeye salmon and has documented substantial time trends in sibling-model parameters, as well as similar trends across groups of sockeye stocks (Holt and Peterman 2004).

Hybrid sibling forecasting model: For some of the Chinook Salmon stocks and age classes examined here, if the fit to a sibling regression model is poor, forecasters tend to use a naïve model (e.g., forecast is the average of the last N years of returns, or perhaps just last year's value). However, the decision of which model to use is not based on any statistical foundation. A more statistically defensible approach is to use the "hybrid sibling" forecasting model developed by Haeseker et al. (2007). "Hybrid" in this case means using a sibling model when the variance of residuals around the relationship is below some threshold (i.e., the data are fit well), but above that threshold using some naïve model. Haeseker et al. (2008) used optimization to find the best variance threshold for each of 40 chum salmon stocks and 37 sockeye salmon stocks. When the performance of the hybrid sibling model was compared to that of 11 structurally different forecasting models using retrospective analysis, it garnered the most sockeye stocks (35%) for which it was the top-ranked model in terms of RMSE and the second-most stocks for chum salmon (29%) (Haeseker et al. 2008).

<u>Generalized forecasting software</u>: ForecastR is a computer program based on the open-source statistical software code, R. It generates age-specific forecasts of salmon abundance (Vélez-Espino et al. 2016). This program is currently in a beta version and is due to be completed by the end of 2016. It is being developed to provide a unified forecasting tool that can be used by researchers and managers across different jurisdictions. ForecastR is flexible enough to be used in different ways in various regions for forecasting abundances of specific stocks. It is intended to facilitate communication and sharing of forecasting results. ForecastR allows users to apply a variety of generic models to their data using various statistical modeling and forecasting tools with the aim to improve the quality of forecasts.

ForecastR will provide a graphical user interface (GUI) to facilitate use by people who do not know how to code in R. The program allows users to forecast abundance of individual stocks (e.g., Chinook, chum, coho, sockeye) based on historical data and other available information. Two types of time series are accommodated, age-specific or total abundances representing individual stocks or aggregates.

Individual analysis modules will provide a variety of capabilities:

- Produce Word or HTML reports (including table of contents, numbered figures and tables with captions, and statistical tutorials)
- Point forecast and bootstrap-based interval forecast
- Numerous diagnostics
- Alternative models (ARIMA, Exponential Smoothing, and Complex Sibling Regressions)
- Probability profiles
- Retrospective evaluations of model performance
- Model ranking, which currently takes place externally from the program

The Panel encourages the further development and application of ForecastR for Chinook Salmon, as long as the program is tested thoroughly first. We suggest the addition of modules for the hybrid sibling and Kalman filter models that are described above.

<u>Dealing with changing parameters</u>: Most fisheries models have one or more components that assume parameters are constant over time, and Chinook forecasting models are no exception. However, extensive evidence exists that parameters such as productivity, marine survival rate, and age-specific maturation rate are "non-stationary", that is, their mean and/or variance has

changed substantially over the years. Non-stationarity thus refers to changes in parameters other than high-frequency year-to-year variability. Non-stationarity affects forecasts of abundance of Chinook Salmon by invalidating the assumptions made in most models that such parameters are constant and not time-varying.

There are at least two approaches to dealing with such non-stationarity. First are methods based on modifying static models to allow estimation of time-varying parameters. The Kalman filter version of the sibling model described above is one example of this approach. It updates parameter estimates annually based on the most recent data and down-weights older data. Truncation of data sets is at the other extreme of methods for dealing with time-varying parameters. Instead of including older data in some parameter estimation step, only data after some cutoff year are used. However, the choice of cutoff year must be made in some defensible manner, such as 1977 in the case of Alaska sockeye salmon because that is when a well-documented "regime shift" occurred to increase productivity of those populations. Arbitrary cutoffs should be avoided.

Sensitivity analyses: Sensitivity analyses are a standard approach to taking uncertainties into account and evaluating their influence on outputs. Sensitivity analyses examine how a given model's output changes with different assumptions or input parameters. Such sensitivity analyses should then be presented as a range of forecasts that reflect model uncertainty. in cases where management objectives have not yet been clarified, sensitivity analyses should also be conducted across the range of plausible model-ranking criteria (e.g., minimizing bias, maximizing precision, etc.).

Multiple models: Agency forecasters should shift their focus away from reporting point estimates of forecasts based on finding the single best forecasting model each year. Instead they should evaluate a set of models and report both a most likely value and the resulting range of point estimates of forecasts along with their respective prediction intervals around the mean forecast abundance. Separately reported forecasts from each model would help to realistically represent some uncertainty in those forecasts. Another option for using information from multiple forecasting models is to combine the forecasts of several of the best-ranking models based on AIC_c weights (which essentially puts a non-zero probability on any one of those models representing the true state of nature and thereby increasing the chance of making a good forecast). This technique of multi-model averaging has a strong theoretical basis (Burnham and Anderson 2002) and is now widely used in ecology (e.g., Connors et al. 2012). A forecast from such a multi-model method would simply be yet another forecast among all the others produced by alternative models.

Centralized relational database: Agency forecasters and the modelers who run the PSC model may obtain increased efficiency in their analyses and production of reports if they were to use a centralized relational database for both input and output data (if they are not already doing so). Such a database can reduce the chance of errors in copying data or formulas in spreadsheets and make creation of new tables and graphs less tedious. The ForecastR program appears to have already addressed the latter issue.

Continuously updated documents: In fisheries, documentation of models, assumptions, and input data is frequently a low-priority task, and such documents often do not get updated, if they are written at all. Such up-to-date documentation is critical, though, both for understanding model results (particularly unusual ones) and for training new people to use the model. A centralized registry or logbook for documentation may assist with this task (section 6.7)

Conclusion D. Establishment of a set of “best forecasting practices" and standard definitions can improve the statistical foundation of methods for stock forecasting.

6.5 Statistical Rigor of Agency Forecasting Methods

Issue V. Forecasting methods for some stocks have not fully incorporated knowledge of changing parameters or recent advancements in statistical methods of analysis.

At the Portland workshop, we learned that past abundance forecasts for the five focal stocks, as well as for the AABMs, have generally been perceived as reasonably sufficient for management purposes. As a result, forecasting methods for both stocks and AABMs have remained largely unchanged from long-standing practices. However, the large forecasting errors in some recent years should create a substantial incentive to explore improvements in those forecasting methods. This section focuses on improvements to stock-specific agency forecasts, whereas section 6.8 below refers to PSC model forecasts.

Most agency forecasts rely heavily on sibling models in which age-specific numbers of fish returning in a given year are projected from historical relationships with numbers of the preceding age class that returned in the preceding year. These models perform best when productivity and maturation rates are stationary (static) over time. However, freshwater and ocean conditions that affect salmon productivity, growth, and maturation are not stable. They vary from year to year and in temporal trends from widely varying patterns in environmental conditions changing at various time scales. Proportions of fish surviving to a given age and maturing in a given year also vary considerably and affect the numbers surviving to, and maturing at, later ages. The Panel learned that present agency forecasting methods do not explicitly incorporate the dynamics of such factors in their analyses. As a result, long-term patterns and annual variations in productivity and maturation rate are likely a key source of forecasting errors. Changes in maturation rate over time are currently addressed in stock-specific forecasting models primarily through trial-and-error fits to data covering different periods to see which assumed period performs best.

Advanced statistical methods may help by taking into account temporal patterns in productivity and/or age-at-maturity schedules that potentially reduce forecasting accuracy and precision. Formal statistical time-series models such as autoregressive lag-1-year (AR1) or autoregressive integrated moving average (ARIMA) models are one promising alternative. As already noted in section 6.4, a Kalman filter estimation procedure may also improve forecasts by taking into account temporal changes in parameters of sibling age-class relationships.

Participants at the Portland workshop identified a variety of other forecasting model refinements with the potential to improve accuracy and precision. Among other things, these included basing age-specific forecasts on all previous ages of the same cohort (e.g., age 4s predicted by the sum of age 2s and 3s), incorporating marine harvest in run reconstructions for estimating terminal run size, and forecasting hatchery and wild fish independently.

Selection of the best model is often based on expert opinions of groups of scientists rather than formal model-selection criteria. It is unknown how different the choice of the best model would be if forecasters used formal, statistically supportable model-selection criteria (Burnham and Anderson 2002) or other approaches such as multi-model averaging. However, at the very least, transparency and confidence in agency forecasts might increase if such well-established formal model-selection criteria were applied. Agency forecasters should also ensure that model-ranking criteria are consistent with management objectives, as noted in section 6.4.

Conclusion E. Accuracy, precision, and transparency of stock forecasting methods might be substantially improved by application of more formal model-selection criteria that match clearly defined management objectives. Forecasts might also improve by use of more advanced statistical methods that allow for time-varying parameters.

6.6 Limitations of Existing Agency Models for Forecasting

Issue VI. Existing forecasting models used by agencies, especially sibling relationships, are reasonably effective in representing average conditions but are vulnerable to performing poorly for years of very low or very high returns.

Years when actual Chinook Salmon abundance is substantially below or above forecast are by far the most challenging for salmon fishery managers. Overestimates of abundance can result in overfishing relative to escapement goals and exploitation-rate/catch limits. Underestimates can result in unnecessary restrictions and substantial foregone harvest. Both situations can substantially disrupt effective fishery implementation and allocation. These errors are particularly troublesome in Chinook Salmon marine fisheries where inseason abundance information is not available to support within-season modifications of fisheries. Forecasting errors are also problematic even in terminal fisheries where inseason information does exist.

The largest discrepancies between actual runs and forecasts typically occur when large numbers of young fish in a cohort do not lead to similarly large numbers of older fish of the same cohort in the next year, or when small numbers of younger fish are followed by uncharacteristically abundant older fish. These patterns are often related to environmentally-driven changes in survival or maturation and are especially common when data used to make forecasts are outside the range of past observations (e.g., abundance of age 3s is larger than previously seen and yet is used to forecast age-4 abundance). Large forecasting errors might also result from a confluence of simple random chance events and measurement error.

Refinements in statistical methods discussed in sections 6.4 and 6.5 may help improve accuracy and precision of forecasts by better accounting for time series patterns in the dataset. However, it is unlikely that more rigorous statistical analysis of existing data will remove the specific problem of large uncertainty in forecasts based on recent observations that are outliers beyond the range of past data. The reason for this prudence in sibling models, for example, is that the width of the prediction interval (frequency distribution of possible forecasts at a given X value) gets wider as abundance increases, and that prediction interval is usually asymmetric (see examples for Columbia River Chinook stocks in section 7).

The best prospect for reducing forecast errors in general, as well as the breadth of uncertainty arising from outliers beyond the range of past observations, will be to develop alternative forecasting methods that explicitly account for dynamic changes in factors that affect salmon returns. Auxiliary information from other independent variables might potentially help account for anomalous return patterns. Examples might include maturation rate in previous years, age-specific body-size distribution, size and age structure of co-varying stocks, freshwater factors, juvenile abundance, juvenile survival indices, or ocean conditions that affect survival or maturation rate. However, with few exceptions (e.g., Orsi et al. 2012), past efforts to incorporate auxiliary information in forecasting models has generally met with only modest success (e.g., Wertheimer et al. 2015). Additional investigations may prove fruitful for Chinook, though. Without new information or new understandings of factors driving variable returns, substantial improvements in forecast accuracy and precision may be difficult to achieve.

We should also recognize that recent poor performance of forecasting methods compared to previous years may be a temporary phenomenon associated with a period of particularly dynamic years for Chinook Salmon abundance throughout the northeastern Pacific. Forecast accuracy and precision may or may not revert to historical norms in the future.

Conclusion F. Development of new models and advanced parameter estimation methods may improve the accuracy and precision of agencies' annual forecasts. Regardless of any such improvements, large uncertainties in forecasts should be expected, especially when they are based on data outside the range of past observations.

6.7 Documentation of the PSC Model's Forecasting Methods

Issue VII. Incomplete and out-of-date documentation of the current PSC Chinook model and its calibration and projection procedures (1) threatens loss of institutional knowledge as key staff move on, (2) increases challenges to new CTC members who want to understand the model and its procedures, and in the worst case, (3) increases the chance of errors in the model's application and interpretation.

Such weak documentation has made it difficult to conduct an effective review of the PSC model's structure and function. For example, a few queries from the Panel about the model based on the available documentation received the reply, "The model doesn't work that way anymore". Follow-up questions did not always help clarify uncertainties. Nevertheless, the Panel is confident about our general conclusions related to the PSC model and its use, in part based on extensive discussions with members of the Analytical Working Group, AWG.

Because of the incomplete documentation of the PSC model, the Panel heard from more than one AWG member that it has taken them up to two years to understand how the model works and even then, further model "experiments" have been needed to determine how outputs are affected by changes in particular inputs. We learned that some parts of the PSC model are only partially understood, even by some who are involved in interpreting the model. Such lack of familiarity with the inner workings and assumptions of the model, as well as the inter-related inputs from the 11 different input files, increases the chance of errors occurring during the model's application and interpretation.

Another aspect of this situation came to the Panel's attention. We learned that one person, John Carlile in Alaska, has the burden (or honor) of taking the lead on annually running the PSC model's calibrations and projections during a short period each spring. Other members of the 12-person AWG subsequently help by running calibrations and making projections of AABM AIs themselves, and checking input and output files for errors. It is unclear to the Panel the extent to which the entire process depends on John's leadership. It is clear, though, that his lengthy experience and intimate knowledge of the model's behavior and code, as well as its error messages, lends critical experience to the PSC modeling process. However, we are unaware of a succession plan (i.e., training of at least one person to take John's place when he moves on). If there is no such plan, we strongly encourage one to be established, particularly in light of the large economic value of Chinook Salmon fisheries that is affected by the PSC model's forecasts of AIs.

Also, apparently there is no single location for registering or logging changes to the PSC model's code, input requirements, or calibration and projection procedures. Such changes may be described in each year's exploitation rate and calibration reports, but the changes are not consolidated in one place, which increases the chance that some subsequent change will unintentionally interact detrimentally with some previous change that is not noticed or remembered. We learned that the AWG has discussed the need for such a central "logbook" of changes to the PSC model, but it does not yet exist. We encourage one to be developed as soon as possible.

Continued application of the PSC Chinook model would benefit substantially from clear and concise up-to-date documentation of its structure, parameter values, assumptions, and data supporting those assumptions.

Conclusion G. Comprehensive up-to-date documentation of the PSC Chinook model in a single, central location is necessary to support its effective and credible use and improvement. A succession plan for training new model users is also critical.

6.8 Statements of Uncertainty about the PSC Model's Output Forecasts

Issue VIII. The deterministic nature of the PSC model and paucity of routine sensitivity analyses do not provide information about uncertainties in the model's forecasts of abundance in the three AABMs and terminal areas, thereby hampering well-informed decision making by PSC Commissioners and fishery managers in AABM areas.

Another element of model transparency deals with documentation of the PSC model's results, i.e., its projections (forecasts) of abundance indices (AIs) for each AABM. Just as CTC members have requested more documentation from regional agency forecasters about the reliability of agency models used for stock-specific forecasting, so too should managers of AABMs benefit from knowing the reliability of past PSC model forecasts of abundances in the AABMs, as well as uncertainties about each current year's forecast AIs.

The Panel learned that each year the CTC provides to the PSC Commissioners and regional managers the forecast abundance index for each AABM and the associated maximum allowable catch (based on Table 1 in Chapter 3 of the Treaty). However, there is no documentation of uncertainties in the annual pre-season forecasts, and no regularly-produced statements or graphs of the long-term performance of the PSC model in terms of comparing pre- and post-season estimates of AIs (Figure 1 above is the only graph we learned about). Only the previous year's point estimate forecast for each AABM is presented annually to Commissioners along with that year's post-season point estimates of AIs.

This omission of measures of uncertainty about stock forecasts unfairly invites criticism when forecast errors inevitably occur. Many assumptions are made by forecasting methods and some, such as constant maturation rate, are not well supported, even by the regions' and the CTC's own data. Given the large number of assumptions made by the PSC model, this lack of statements about uncertainty in forecasts can mislead decision makers about the real situation and can lead to inappropriate regulatory decisions. Omission of such information also keeps managers from making well-informed decisions in which risks (created by uncertainties) are traded off against potential benefits (also uncertain). Other, more complex stock assessment models on many other fish species routinely produce measures of uncertainty around forecasts for fisheries managers (National Research Council, U.S.A. 1998; Walters and Martell 2004; Fisheries and Oceans Canada 2015), so it is not unreasonable to expect it from the PSC model as well. PSC modelers should incorporate the routine practice of showing uncertainties around their point estimates of annual abundance indices (AIs) in each the three AABMs.

The advice about regularly conducting extensive sensitivity analyses given above in section 6.2 for agency stock-specific forecasts is equally important for the PSC model. Sensitivity analysis of model output to different assumptions and input data is a standard means of taking uncertainties into account and propagating their effects all the way to model outputs. However, at the workshop the Panel heard that the annual calibration of the current PSC model is too time consuming to conduct more than a few such sensitivity analyses each year. From the Panel's perspective, this response highlights a key limitation of the current PSC model. Given the numerous uncertainties that are known to exist in the real-world system represented by the model, the inability to conduct large numbers of sensitivity analyses (e.g., dozens or more) severely constrains the level of confidence that can be placed in results from the PSC model (not just forecasts of abundance in the AABMs, but also comparisons of pre- and post-abundance estimates).

The Panel recognizes the severe problem of limited CTC staff time (especially for members of the AWG). Staff face competing demands to provide a complex set of annual reports during a short period for implementing the Pacific Salmon Treaty, testing and documenting existing methods and models, and exploring improvements in related methods of analysis. Difficult decisions will obviously have to be made by the AWG to rank tasks when all are important but all cannot be thoroughly addressed with existing staffing. One way out of the dilemma of too much work and too little time is to reduce the extent of the CTC's reporting requirements to the PSC by negotiating changes to the next Treaty. Another is to conduct more analyses outside of the intensive early-spring period.

More detailed discussion and recommendations for addressing uncertainty may be found in sections 7 and 8 of this report.

Conclusion H. Point estimates of forecasts of abundance indices in the three AABM areas from the PSC model should be accompanied by descriptions of uncertainties in those forecasts. Uncertainties can be derived from extensive sensitivity analyses of effects of different assumptions and input parameters. Expression of uncertainty in these forecasts is essential for determining the confidence to be placed in them and allowing for appropriate consideration by fishery managers.

6.9 Limitations of the Existing PSC Chinook Model

Issue IX. The PSC model's structure, parameterization, and calibration are complex and subject to substantial structural and parameter uncertainties.

Objectives and applications of the PSC Chinook model have evolved over time. The model appears to serve a number of purposes, some for which it may not be ideally configured. These include calculation of an abundance index in currency similar to that used in the base period to establish harvest control rules, inferences of abundance and harvest for stocks for which assessments are not independently derived, estimation of fishery impact rates for reference to objectives, and projections of terminal run sizes in some fishery areas. Estimation of the abundance index in the three AABM areas appears to be the most important current application. The ranking of future PSC modeling tasks could benefit from more explicit definitions of priority uses.

The current model structure is relatively unchanged since 1985. During stage-1 calibration, it assumes that all factors are known except one, the time series of stock-specific EVs, which is estimated by fitting the model to the input data. Trends and variability in productivity and maturation rate can lead to uncertainty in the forecasts of AIs. The PSC model currently attempts to take such changes into account by using base-period maturation rate along with annual CWT data, the most recent 9-year average maturity rate (but only for dealing with incomplete broods), and fitting a time series of EV values to observed abundances. The resulting uncertainties in forecasts of abundance indices in AABMs are not clearly articulated in the CTC's annual reports.

Previous sections have already described the need for additional testing and refinement of the PSC Chinook model. The Panel also recognizes that opportunities for such work are limited by competing work demands on members of the CTC. These demands appear to be hampering exploration of alternative and possibly improved forecasting methods.

Given this predicament, continuing exploration of other modeling options is appropriate. Significant effort has already been invested in developing, and in some cases applying, alternative fisheries models. Details of some of these options are provided in section 8.

Conclusion I. Substantial revision, testing, or possibly even replacement of the existing PSC Chinook model is necessary to effectively serve continuing needs, including the need for statements of uncertainty in the model's forecasts. A subgroup of CTC members should be created to explore such revisions and new models.

6.10 Consistency of Management Structures/Policies with the Limitations of Information and Assessments

Issue X. Limitations of data and uncertainties associated with stock assessments and forecasting models challenge effective implementation of abundance-based management of Chinook under the Pacific Salmon Treaty.

Implementation of an effective abundance-based management strategy, as specified in the Treaty, obviously requires sound abundance forecasts and information on harvest rates. Forecasting problems in recent years have hindered effective implementation of abundance-based management in AABM fisheries when post-season abundance estimates have differed greatly from pre-season forecasts.

One avenue for addressing these problems is to improve forecasts by collecting better data, modifying current models or building better ones, and reporting uncertainties about forecasts and post-season estimates. These topics are the main focus of this report, which includes a variety of related recommendations.

Scientists are often inclined to pursue ever more-detailed data and fine-scaled models in an attempt to explain and reduce uncertainty. The danger in this approach is that our expectations and models can easily outstrip the fundamental limitations of the available information and resources. Thus, the other avenue for dealing with forecasting problems is to scale expectations for the forecasts and management strategies to match the limitations of the existing information and methods of analysis. More complex, finer-scale, mechanistic models are not always a better answer. Simpler, more transparent assessments and strategies often prove every bit as effective in achieving desired outcomes as more complex but subjective models loaded with assumptions. The Panel recognizes that references to base-period abundances, age structures, and exploitation patterns are important features of the current Chinook modeling structure, but a more streamlined framework could also be configured to do so.

The existing limitations of CTC reporting, PSC model documentation, testing, and refinement suggest to the Panel that the current analytical framework for Chinook management under the treaty warrants a close look to determine whether an alternative process might be more appropriate. We discuss such alternatives in section 8.

Conclusion J. Alternative frameworks, as well as ways of using forecasts of abundance, should be considered for Chinook if current information and resources are not sufficient to effectively conduct adequate analyses and implement provisions of the current Treaty. Those provisions may need to be changed during current negotiations.

7 Regional Agency Forecasts of Chinook Abundance

7.1 General Comments about Agency Forecasts

This section focuses on the five Chinook Salmon stocks investigated as part of this review. Regional agencies generate forecasts for those stocks and send them to the CTC for input to the PSC's coast-wide model. In addition to these forecasts, some regional agencies such as DFO also produce other forecasts for use in management of their domestic or terminal fisheries. This section only discusses the first type of forecasts, the ones used as inputs to the PSC model. The agencies' forecasts are generally sent to PSC modelers by March 1st of each year, but the domestic forecast for WCVI, for instance, is generally done later in the spring when more information from the previous fishing season is available.

Before we deal with individual Chinook Salmon stocks, we cover several topics and recommendations that are relevant to agencies' forecasts for all five stocks.

First, as mentioned in section 6.2, the Panel learned that there is very little communication between the CTC and biologists in the regions who annually submit their stock-specific forecasts to the CTC for input to the PSC model. The CTC thus has no information on the model-ranking process or the reliability of those forecasts.

Recommendation 1. ***When regional agency forecasters send their stock-specific annual forecasts to the CTC, they should document their model-ranking procedures as well as the past performance of their methods (bias and precision).***

Second, a key repeating theme in our review of Chinook Salmon forecasting methods is the limited representation of uncertainties in both analyses and the resulting forecasts. The types of uncertainties described in section 5 apply to all Chinook Salmon stocks and need to be explicitly considered when making forecasts. Omission of such uncertainties creates overconfidence in forecasts and may lead to inappropriate management regulations and outcomes. Details about the importance of considering and reporting uncertainties, and methods for doing so, have already been provided in sections 6.2, 6.4, and 6.5, so we refer readers back to those sections. Suffice it to say here that regional agency forecasters can go further than at present toward developing a range of forecasts that reflect various types of uncertainties in their analyses.

Uncertainties about how to represent the natural system are unavoidable when choosing the structural form and parameter values of forecasting models. It is therefore important that forecasters explicitly recognize those uncertainties during their analyses. One way to do so is to admit that the "best" single model, however that is determined, does not have a probability of 1.0 of being the correct representation of nature, and that other models might be useful to consider as well when describing uncertainty in forecasts. Section 6.4 elaborated on this need to avoid the current agency practice of focusing on the single best model.

The set of forecasting models that each agency should consider could be defined using formal model-selection methods (Burnham and Anderson 2002). For example, that set could be identified as those having delta AIC_c values less than some number (for example, 4 or 6). The resulting set of models would be those that have sufficient support to be considered plausible descriptions of the natural system (Burnham and Anderson 2002). The range of forecasts of abundance produced by this set of models would then reflect structural uncertainty and some aspects of parametric uncertainty.

To give a more complete picture of these uncertainties, point forecasts from each model should be accompanied by prediction intervals (not confidence intervals). A prediction interval illustrates the probability distribution from which the single point estimate of the forecast is drawn. For salmon forecasting models, prediction intervals are typically quite large (e.g., Figure 17 for sibling models for three Columbia River Chinook Salmon stocks). These prediction intervals should therefore be passed on to the PSC modelers (CTC) so that the resulting uncertainties can be reflected in the PSC model's forecasts of abundance in the AABM areas. Prediction intervals may be more important than structural uncertainty in terms of generating uncertainty about future abundances, although this may depend on the particular stock and data set. Only future analyses can tell whether the structural uncertainty arising from structurally different models is important enough to justify sending separate forecasts of each model to the CTC for use in the PSC model, but regardless, prediction intervals should be sent to PSC modelers.

Recommendation 2. ***Agency forecasters should not choose just one best model for forecasting abundance in each age class. Instead, they should conduct analyses across different models that make different assumptions and report the resulting set of forecasts to the CTC for use as inputs to the PSC model. The generally large prediction intervals (not confidence intervals) around point forecasts should also be reported.***

Third, it is often the case in fisheries that the rank order of forecasting models can be substantially affected by which model-ranking criteria are used. The model that has the smallest bias (MPE closest to zero) is often different from the model that has the greatest precision (lowest MAPE). A model that does well with both attributes (reflected by the lowest RMSE) may differ yet again. Model forecasts in a given year can also differ between such models. It is therefore important that forecasters carefully choose their model-ranking criteria. As elaborated upon in section 6.4, the chosen criteria should be consistent with the way that management objectives are stated. Management concerns about long-term statistical bias of forecasting methods would be addressed by using MPE as the model-ranking criterion, whereas MAPE would be appropriate if the greatest concern is the magnitude of yearly deviation between pre- and post-season estimates.

The Panel heard at the Portland workshop, as well as afterwards, that there is uncertainty among agency forecasters about management objectives, both at the regional level and for AABMs (Appendix C). Such clarity is essential so that agency forecasters can design their model-ranking criteria to produce the most useful information. Until such time as those management objectives are clearly articulated, the Panel recommends that agency forecasters produce forecasts for the top model for each of the three most common ranking criteria, MPE, MAPE, and RMSE. After clear objectives are stated, then only forecasts from the models using the appropriate ranking criteria need to be reported. In all cases, forecasts should be accompanied by their respective prediction intervals, as described in the previous recommendation.

Recommendation 3. ***Agency forecasters should also send to the CTC a set of forecasts, each one based on a different model-ranking criterion, as determined by stated management objectives. As described in section 8.2, the CTC can then conduct sensitivity analyses with the PSC model to determine their effect on forecasts of abundance in the AABMs.***

Fourth, the ForecastR software program explained above in section 6.4 has the potential to improve the statistical basis for stock forecasts in all regions. From what the Panel has seen of the output from a beta version of ForecastR for Northern Oregon Coast Chinook Salmon, this program promises to be a very useful, standardized method for regional agencies to produce their forecasts. It allows users to choose from a wide variety of forecasting models.

Recommendation 4. ***We encourage all agency forecasters to try applying ForecastR to their regions' stocks. As well, the CTC should run workshops to familiarize agency scientists with the ForecastR program.***

Fifth, sibling regressions do well when the residual variance around the line is small. When this is not the case, forecasters often revert to some type of naïve, or non-biologically based model such as simply using last year's abundance or an average of abundances over several past years. As described in section 6.4, a "hybrid sibling" model provides a statistically sound basis for choosing either a sibling model or a naïve model in any given year (Haeseker et al. 2007). This hybrid model performed well in retrospective analyses.

Recommendation 5. ***Agency forecasters should try applying a hybrid sibling model, especially to cases in which the fit of data to a standard sibling model is weak.***

Sixth, the CTC has documented a decrease over time in mean age-at-maturity among several west-coast Chinook Salmon populations, as well as changes in survival rates (CTC 2016a). These two changes therefore violate a key assumption underlying the fitting of sibling age-class relationships, namely that parameters are constant over time. Such non-stationarity undoubtedly contributes to forecasting errors, but the Panel does not have enough information to state the magnitude of that contribution relative to other sources of forecasting error. Essentially, such changing maturation rate and/or survival rate would result in time-varying slope and/or intercept of a sibling relationship. A Kalman filter estimation procedure for the sibling relationship estimates such time-varying parameters (Holt and Peterman 2004; also section 6.4). Although Holt and Peterman (2004) used a random-walk residual term in the system equation, an AR1 term may work even better if the changes in maturation rate and/or survival rate are highly autocorrelated in time.

Recommendation 6. ***We recommend that agency forecasters try using a Kalman filter estimation procedure for fitting their sibling relationships to account for time-varying parameters.***

Seventh, all five of the agency forecasts reviewed were missing a complete assessment of stock contributions to the various AABM and ISBM fisheries. Instead, in all cases it was left to the PSC modelers to fill in the missing components. Although the Panel does not fully understand how those estimates are made, we recognize that surrogate groups of fish marked with coded-wire-tags have to be used to represent larger stock groups. For the vast majority of these cases, it is well known that because the marked group is a small component of the overall stock it represents, it may not mimic the actual exploitation and survival rates of the larger stock. It was demonstrated that forecasts of terminal return and escapement can be made without fishery contribution estimates, but the quality of forecasts might improve with better estimates of age composition and numbers of fish in both harvest and escapement. These are the building blocks of strong population assessment programs, as well as good forecasting.

The Panel learned that the United States and Canada both contributed $7.5 million to their agencies in recent years to improve their coded-wire-tag programs. We applaud this undertaking but also worry about the form the program will take in upcoming years without supplemental funding.

While there have been dramatic improvements in estimating stock compositions using CWT's, there still are problems with the ability to make those estimates timely. Final estimates from fisheries are typically not available for up to two years after the tags are recovered. Methods to reduce the time delay and improve the usefulness of the information should be evaluated. An example would be the use of in-season creel surveys to obtain marked-to-unmarked fractions prior to receiving the results from post-season mail-out surveys of anglers.

Otolith mass marking has replaced CWT's in many fisheries for estimating stock contributions (Hargreaves et al. 2001; Joyce and Evans 2000). Because otolith marking allows for marking greater numbers of fish, the estimation of stock contributions is often more efficient and timely than with CWT's. The downside to the use of otoliths is the low number of different discernable patterns available to be applied to the otoliths which ultimately limits the number of stocks that can be distinguished.

Genetic methods have found increasing applications in the management of Pacific salmon over the past 20 years (Beacham et al. 2008; Gilk-Baumer et al. 2013). Technologies for the application of these methods have improved rapidly, and it has become increasingly feasible to collect and process large quantities of genetic data in a timely manner at reasonable cost. It is reasonable to expect these new technologies will continue to evolve and become more valuable for the management of ocean salmon fisheries (Pacific Salmon Commission 2008). While there are other methods to explore, parentage-based tagging is a genetic technique that shows promise as an efficient alternative to physical tagging methods such as coded-wire-tags (Larson 2014; Anderson and Garza 2006).

Recommendation 7. ***Continue to improve upon the ability to estimate the contribution by stock to all AABM and ISBM fisheries with the objective of obtaining reliable stock contribution estimates by age. The Panel encourages the commitment of extra funding for the implementation of techniques to estimate stock contributions in a timely enough manner that the results can be used for forecasting in the subsequent year.***

7.2 Columbia River

7.2.1 Forecasting Model

The Panel reviewed forecasts for three stocks of Chinook Salmon from the Columbia River -- Upriver Brights, Spring Creek Hatchery, and the Summer run. The fisheries in the Columbia River are managed subject to provisions of the continuing jurisdiction of the Federal court in proceedings between the United States, and the State Agencies and Treaty Tribes. A Technical Advisory Committee (TAC) consisting of representatives from each of these entities is tasked with assessing and forecasting salmon returns to the Columbia River. The methods for forecasting all three Columbia River stocks are very similar, so the Panel's comments here apply generally to all three stocks.

A large number of relationships (more than 80 each for the Upriver Bright and Spring Creek Hatchery stocks alone) are examined each year for potential use in forecasting. These relationships include a combination of sibling regressions, cohort ratios, and the average of recent returns. The large number of relationships used to forecast a particular age group arises from exploring numerous sets of brood years of data that are chosen as input to the analyses. The TAC uses an open discussion among its expert panel members to select a relationship for each of the age classes within a stock to use in making the final forecast. The selection criteria include r^2 values, forecast bias (mean percent error), and forecast precision (mean absolute percent error). Sibling regressions have been selected most of the time (Figure 16). Years to include in the regressions were selected based on recent forecast performance. An average of recent-year returns is occasionally used when a particular age component represents a minor component of the entire return (e.g., <1,000 fish).

While it could be argued that the acquired knowledge of the TAC's experts is sufficient for producing accurate forecasts, people occasionally leave and others move into the group, resulting in both a loss and introduction of knowledge and experience. A more formalized, statistically based process for selecting the final forecasting relationships might lead to improved forecasts. As a starting point, we suggest <u>not</u> using r^2 as a model-ranking criterion, but instead using the more directly relevant measures of forecasting performance: MPE, MAPE, or RMSE. The Akaike Information Criterion (AIC_c) discussed in Burnham and Anderson (2002) may also be useful, <u>but with the caveat mentioned at the end of Example 2 in section 8.2</u>.

Recommendation 8. ***The Columbia River Technical Advisory Committee (TAC) should explore whether using formal statistical model-selection criteria improves the accuracy and precision of their forecasts.***

All of the sibling relationships examined for the Columbia River Upriver Bright and Spring Creek Hatchery stocks were regressions using the untransformed number of Chinook Salmon by age estimated to have returned to the terminal area, while some of the relationships used to forecast Columbia River Summer stock were $\log_e$-$\log_e$ regressions. Peterman (1982) demonstrated that the use of natural log transformations of sibling data provided relationships that better meet the assumptions of regression such as constant variance across the range of X values. The presentations made by Steve Haeseker and Stuart Ellis at the workshop indicated that the regression fits to the untransformed and transformed data were similar for the majority of observed terminal run sizes, but there was a slight divergence at larger run sizes (Figure 17). While it is easier to explain a regression to non-technical audiences using untransformed data, it may be advantageous to explore the use of natural-log transforms for forecasting. This may be especially appropriate now because recent run sizes are at the upper end of the historically observed data.

Recommendation 9. ***Explore the use of natural-log transformations for sibling regressions. The examination should evaluate both the effect on meeting the regression assumptions and forecasting performance.***

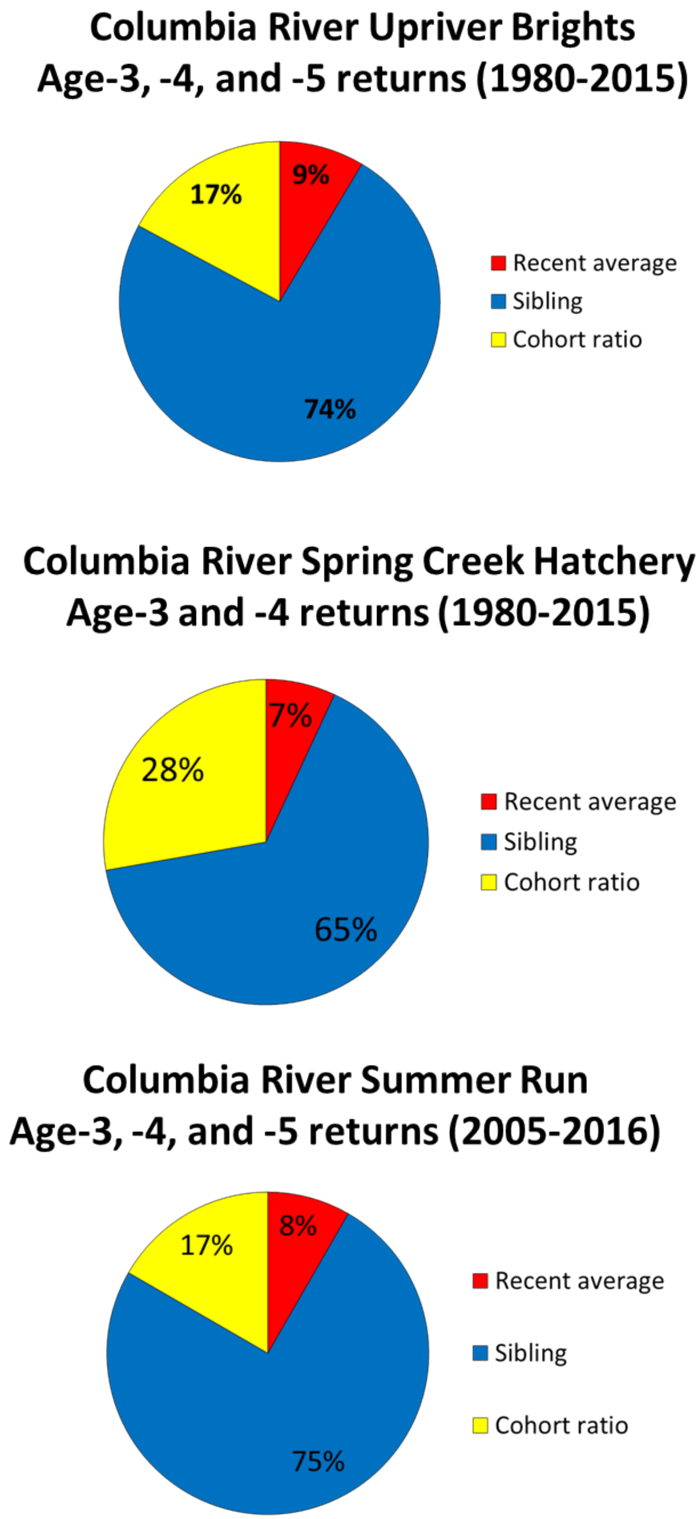
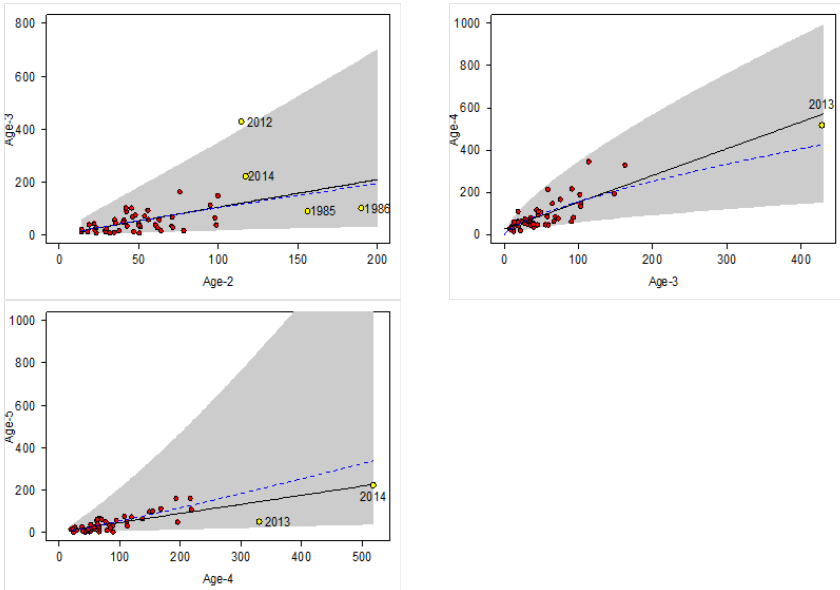
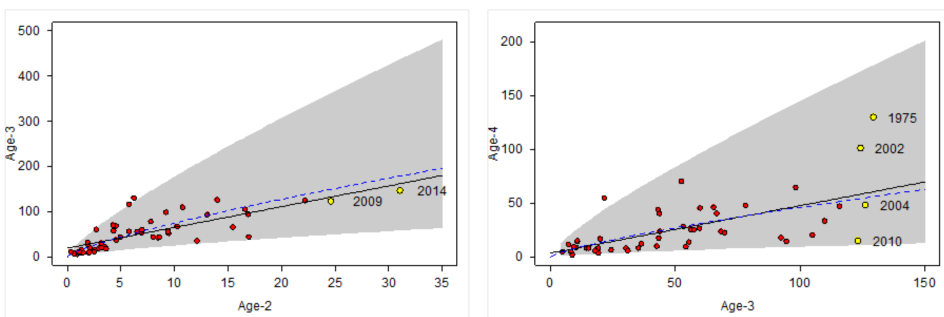


***Figure 16.** The distribution of models selected in the past by the Columbia River Technical Advisory Committee (TAC) to forecast Columbia Upriver Bright, Columbia River Spring Creek Hatchery, and the Columbia River Summer stocks of Chinook Salmon. From the information presented by Steve Haeseker (USFWS) and Stuart Ellis (CRITFC) at the Portland workshop, August 10-11, 2016.*

Upriver Bright sibling relationships (BY 1962-2012)



Spring Creek sibling relationships (BY 1962-2012)



Summer Chinook sibling relationships (BY 1986-2012)

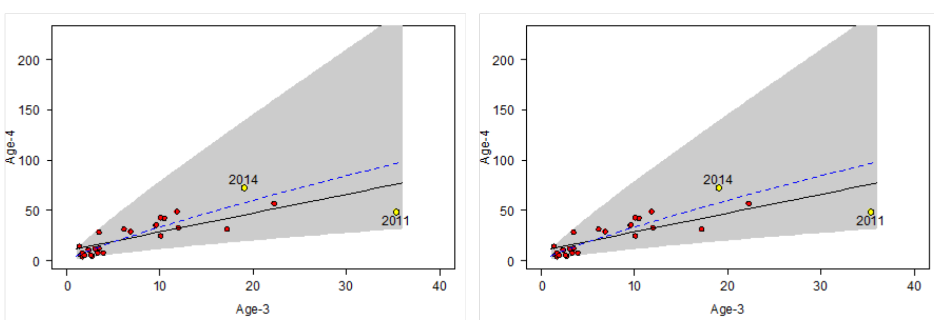


***Figure 17.** Sibling relationships for forecasting Columbia River Upriver Bright, Columbia River Spring Creek Hatchery, and Columbia River Summer stocks of Chinook Salmon. The solid black line indicates the fit of untransformed salmon numbers while the dashed blue line indicates the fit of natural-logged data. Open-circle data points indicate recent extreme abundances. From the information presented by Steve Haeseker (USFWS) and Stuart Ellis (CRITFC) at the Portland workshop, August 10-11, 2016.*

Inherent to both the sibling regression and average ratio methods for forecasting is the assumption that each age class contributes a constant proportion to their brood-year total returns. In other words, it is assumed that there is no year-to-year variability in that proportion and no long-term trend. However, it is clear that the proportion of each age in the brood year returns of all three Columbia River Chinook Salmon stocks examined is not constant (Figure 18, Figure 19, and Figure 20). While some of the variability in the proportion of age in the brood-year return is natural variability in the maturation and/or survival rates, it is more than likely that a portion of the variability is due to changes in exploitation rate between years in the AABM and ISBM fisheries.

The PSC model annually forecasts Abundance Indices (AIs) each year for the three AABM fisheries, which are used to set maximum allowable harvests in those areas. Because the AI in a given area is dependent upon the magnitude of up to as many as 28 stocks in the PSC model, it varies from year to year, and ultimately results in varying exploitation rates from year to year. This is especially true when an AI moves between tiers of harvest rates (Table 1 of Chapter 3 of the Treaty). In addition, migratory pathways of salmon stocks are seldom consistent from year to year, and it is likely that stocks experience different rates of exploitation solely based on how and when they enter the fishing areas and how long they remain. As noted near the start of section 7, obtaining estimates of both the contribution of Columbia River Upriver Brights, Columbia River Spring Creek Hatchery, and Columbia River Summer stocks by age to these fisheries and the corresponding estimates of uncertainty about those estimates should provide additional information for dealing with the variability in age composition due to the unequal exploitation.

7.2.2 Forecasting Performance

The forecast of Columbia Upriver Bright Chinook Salmon provided for use in the Pacific Salmon Commission (PSC) model is nearly unbiased and reasonably precise (mean percent error = 1%; mean absolute percent error = 25%; calculation done by Brian Bue using data from CTC 2015b Appendix J1 for 1999-2013 and preliminary data for 2014 and 2015 obtained from John Carlile, ADF&G). The most recent years (2013-15) were the largest returns in the data set and showed the greatest deviation between the forecast and actual terminal returns (Figure 21). There was no obvious time trend of forecasts being either under- or overestimates (Figure 22).

The forecast of the Columbia River Spring Creek Hatchery stock provided for use in the Pacific Salmon Commission (PSC) model shows a tendency to overestimate the actual abundance and is reasonably precise (mean percent error = 8%; mean absolute percent error = 31%; Figure 23, calculation done by Brian Bue using data from CTC 2015b Appendix J1 for 1999-2013 and preliminary data for 2014 and 2015 obtained from John Carlile, ADF&G). There are periods of underestimates of the run (2001-2004) and overestimates (2005-2012; Figure 24).

Upriver Bright age-class proportions (BY 1962-2010)

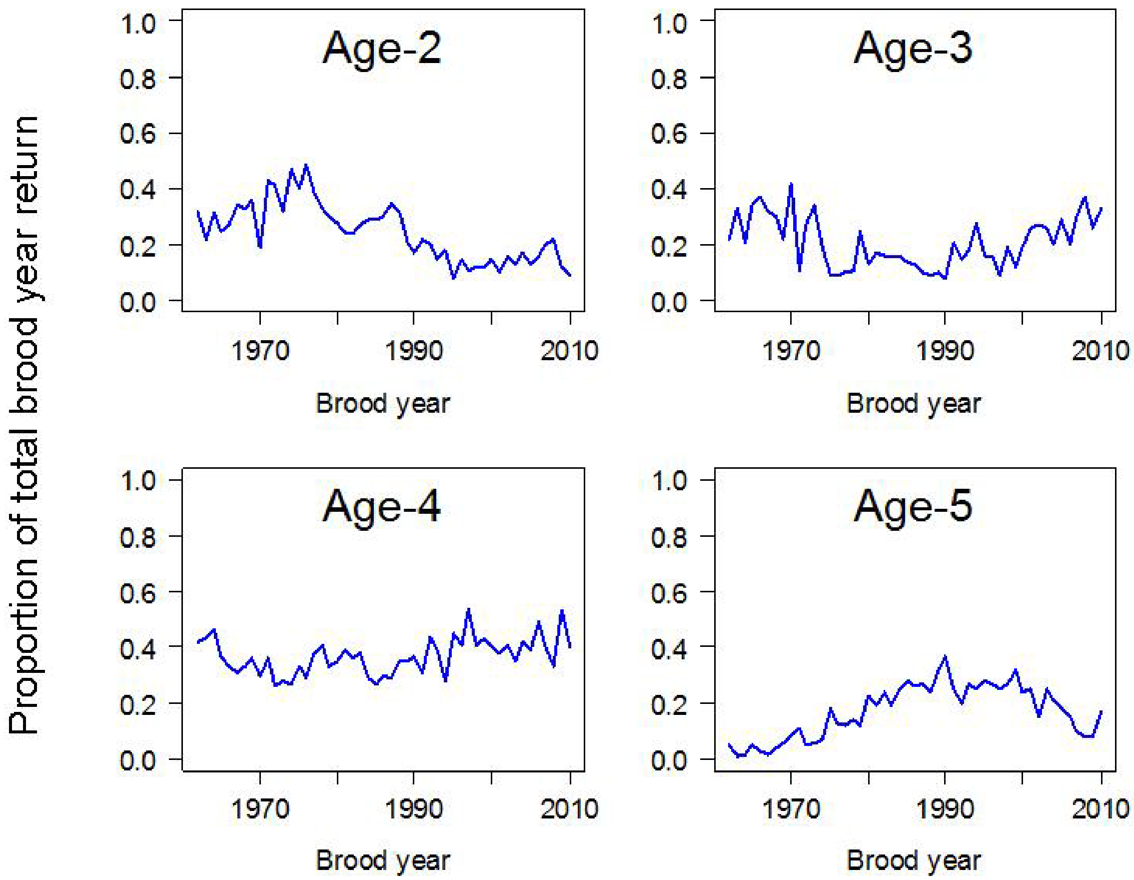


Figure 18. The proportion of total brood year return by age class and year for Columbia River Upriver Bright Chinook Salmon, 1962-2010. From personal communication by Geoffrey Whisler (ODFW), presented by Steve Haeseker (USFWS).

Spring Creek age-class proportions (BY 1962-2010)

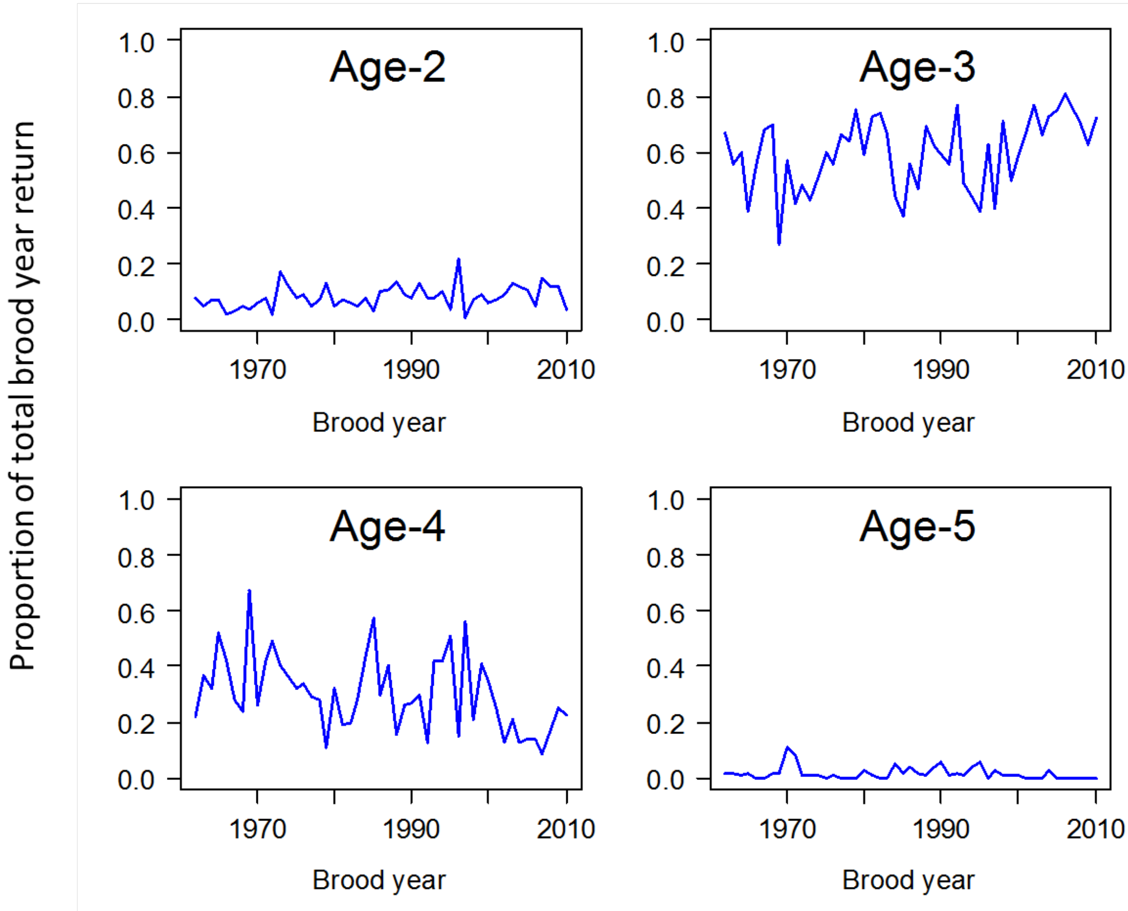


Figure 19. Proportion of the total brood year return by age class and year for Columbia River Spring Creek Hatchery Chinook Salmon, 1962-2010. From personal communication by Geoffrey Whisler (ODFW), presented by Steve Haeseker (USFWS).

Summer Chinook age-class proportions (BY 1986-2009)

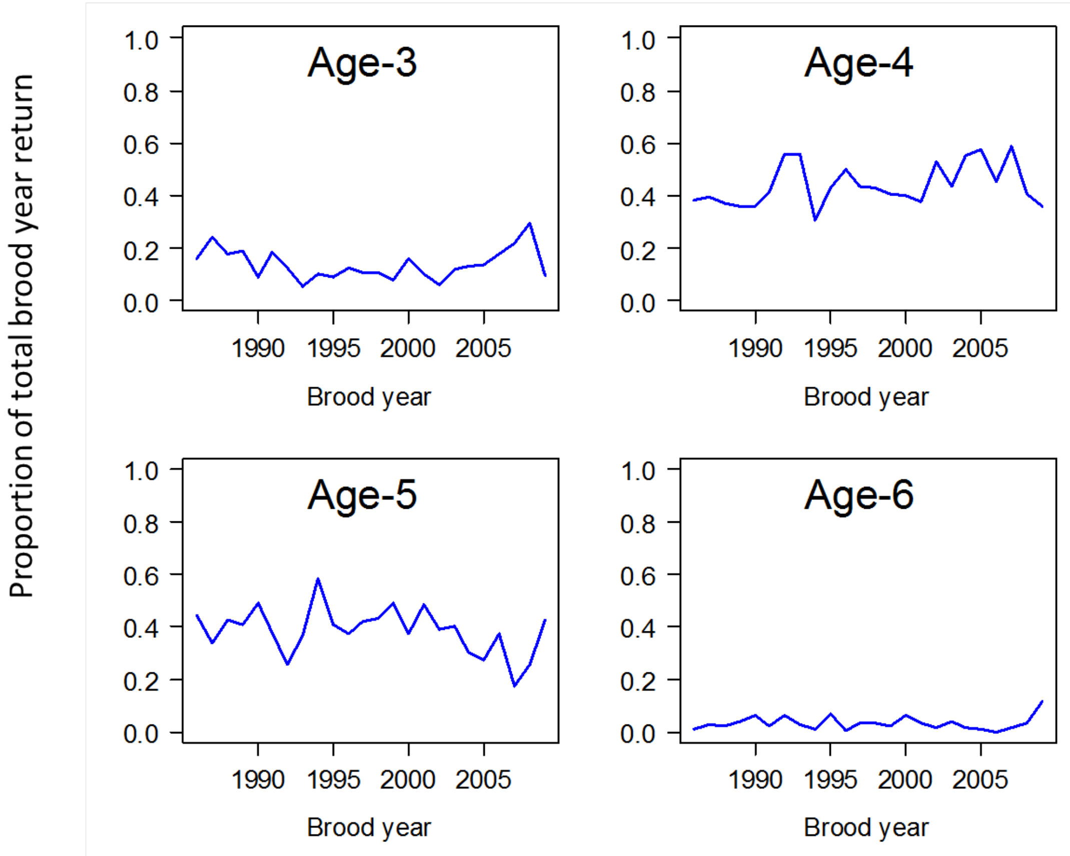


***Figure 20.** Proportion of total brood year return by age class and year for Columbia River Summers, 1986-2009. From personal communication by Geoffrey Whisler (ODFW), presented by Stuart Ellis (CRITFC).*

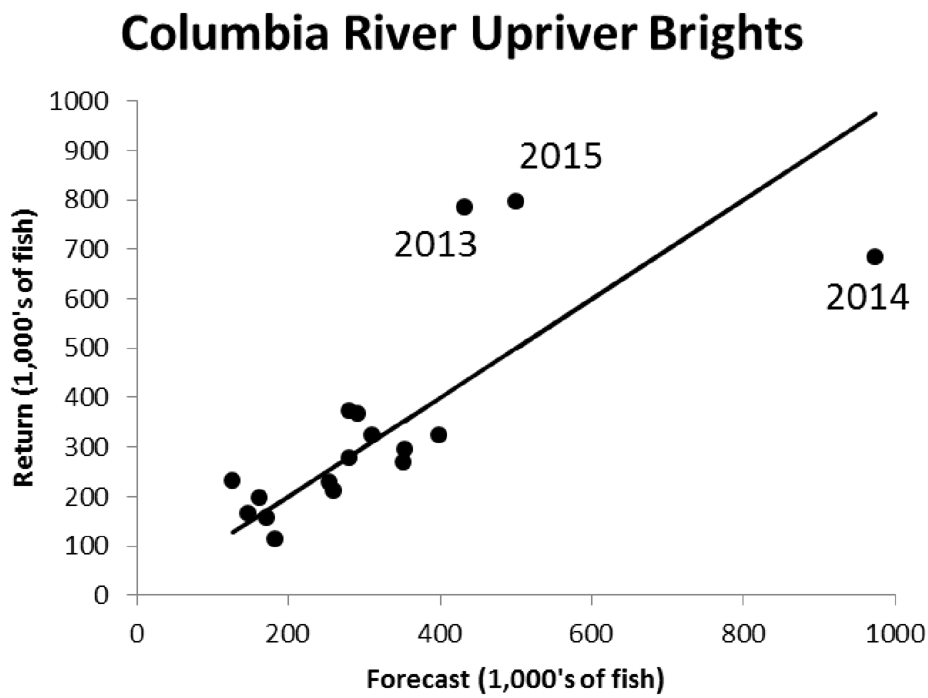


Figure 21. Comparison of the preseason terminal forecast of Columbia River Upriver Brights provided for use in the PSC model with the actual terminal return, 1999-2015. 1999-2013 data from CTC 2015b Appendix J1; 2014 and 2015 are preliminary and obtained from John Carlile ADF&G).

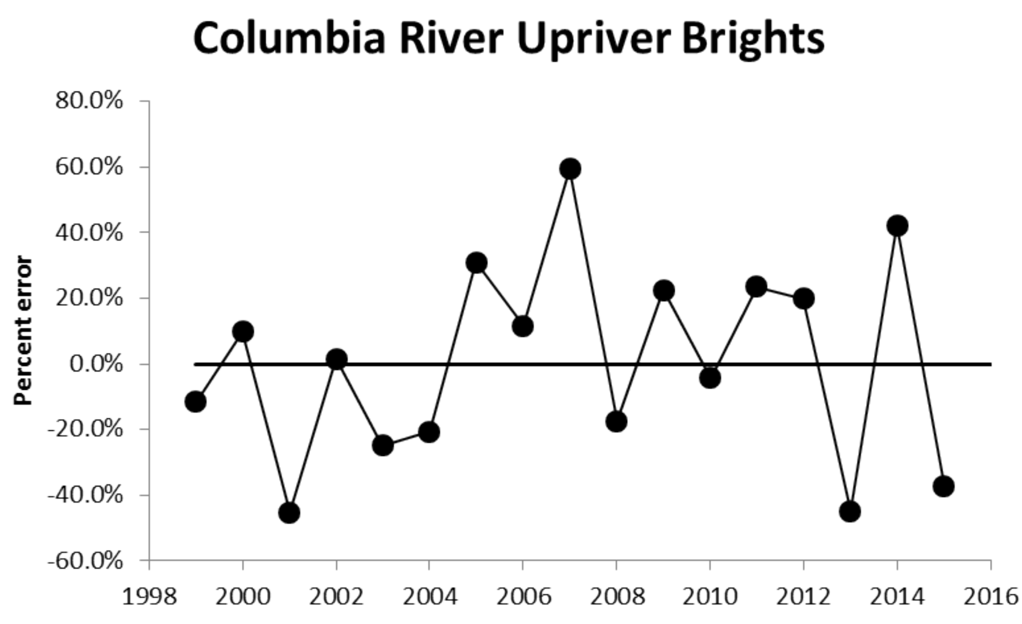


Figure 22. Percent error by year for the forecast of the Columbia River Upriver Bright stock provided for use in the PSC model, 1999-2015. Percent error was calculated by Brian Bue as (forecast – return)/return; positive values indicate the forecast was larger than the return. The 1999-2013 data from CTC 2015b Appendix J1; 2014 and 2015 are preliminary and obtained from John Carlile ADF&G).

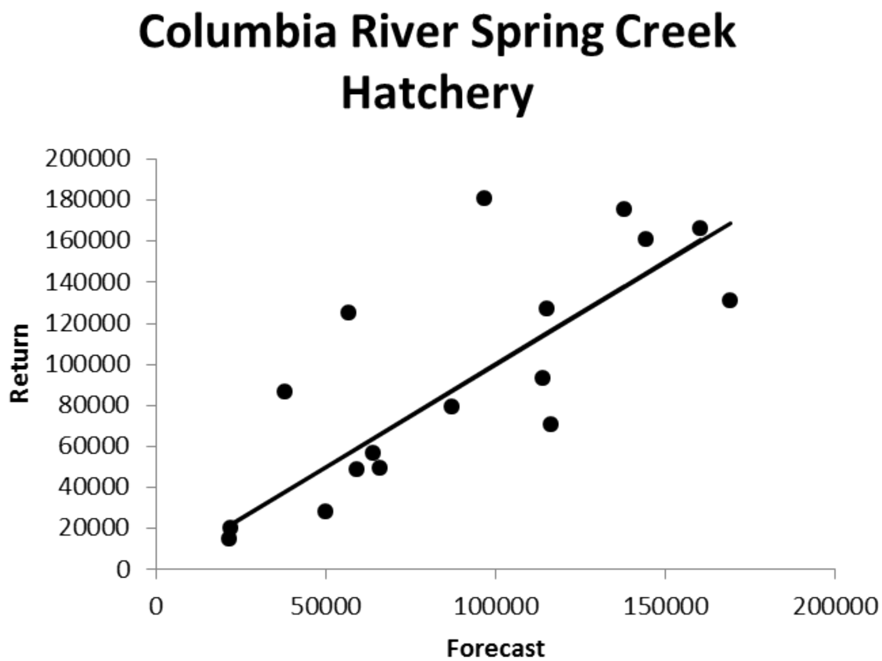


***Figure 23.** Comparison of the preseason terminal forecast of Columbia River Spring Creek Hatchery provided for use in the PSC model with the actual terminal return, 1999-2013. 1999- 2013 data from CTC 2015b Appendix J1; 2014 and 2015 are preliminary and obtained from John Carlile ADF&G).*

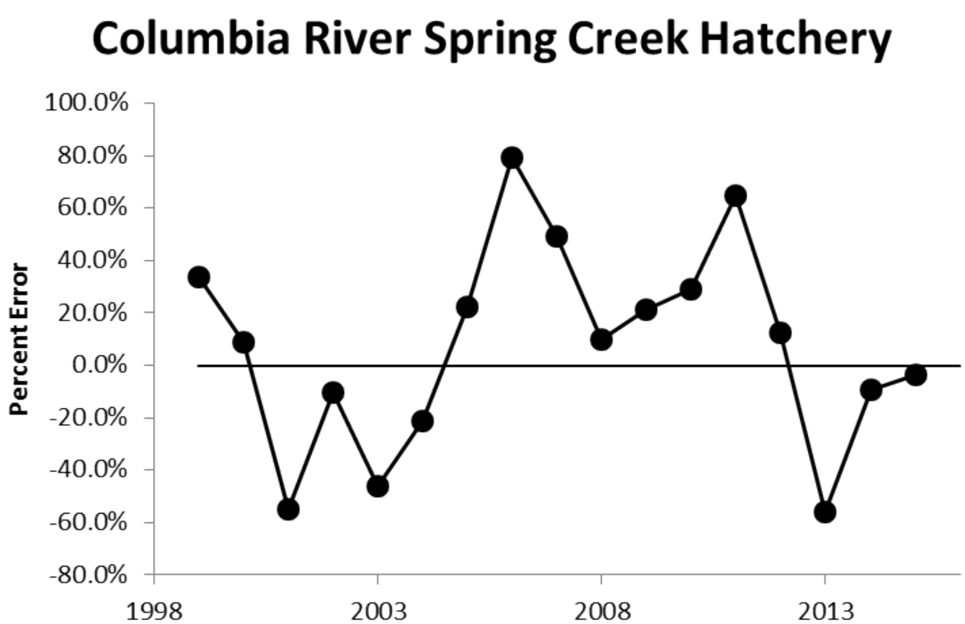


***Figure 24.** Percent error by year for the forecast of the Columbia River Spring Creek Hatchery stock provided for use in the PSC model, 1999-2013. Percent error was calculated by Brian Bue as (forecast – return)/return; positive values indicate the forecast was larger than the return. The 1999- 2013 data from CTC 2015b Appendix; 2014 and 2015 are preliminary and obtained from John Carlile ADF&G).*

Although forecasts of the Columbia River Summer run have been made since 1999 for input to the PSC model, it is only appropriate to compare the forecasts and agency estimates of actual return for the years 2005 to the present. Prior to 2005, the Columbia River Summer stock was defined to be those fish that entered the river between June 1 and July 31. In 2005, the definition of the Summer stock was changed to those fish that entered the river between June 15 and July 31.

Recommendation 10. ***The Columbia River Technical Advisory Committee and the Pacific Salmon Commission's modeling group should communicate with each other to ensure that they are both working with the same definition of the Columbia River Summer stock and the same sets of data, and that any historical information reflects this change.***

The forecast of Columbia Upriver Summer Chinook Salmon provided for use in the Pacific Salmon Commission (PSC) model for the years 2005-2015 shows a tendency to overestimate abundance and is reasonably precise (mean percent error = 10%; mean absolute percent error = 22%). The forecasts for 7 of the 8 years forecast between 2005 and 2013 were greater than the observed returns, whereas the returns for 2014 and the record-high 2015 were more than forecast (Figure 25, Figure 26).

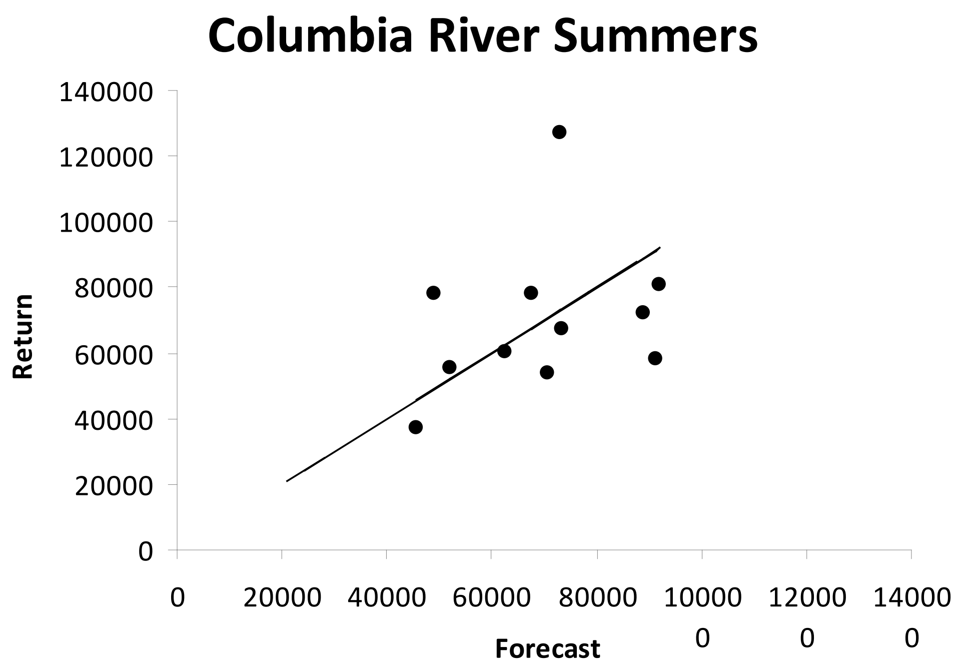


***Figure 25.** Comparison of the preseason terminal forecast of Columbia River Summers provided for use in the PSC model with the actual terminal return, 2005-2013. The 2005-2013 data from CTC 2015b Appendix J1, except 2006 from Jeff Whisler, ODFW; 2014 and 2015 are preliminary and obtained from John Carlile ADF&G).*

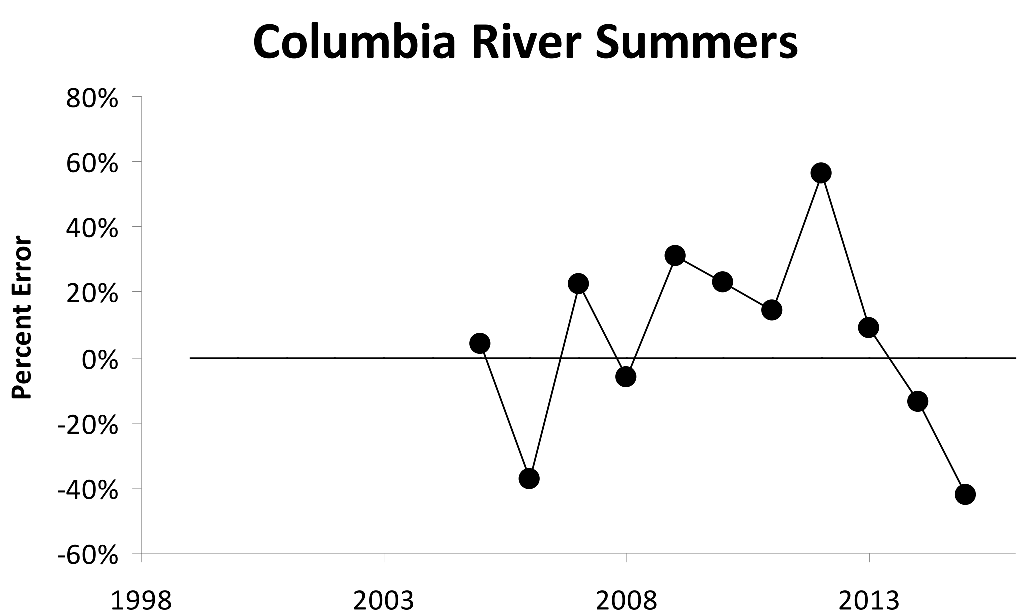


***Figure 26.** Percent error by year for the forecast of the Columbia River Summer stock provided for use in the PSC model, 1999-2013. Percent error was calculated by Brian Bue as (forecast – return)/return; positive values indicate the forecast was larger than the return. The 2005-2013 data from CTC 2015b Appendix J1, except 2006 from Jeff Whisler, ODFW; 2014 and 2015 are preliminary and obtained from John Carlile ADF&G).*

7.3 West Coast Vancouver Island

7.3.1 Forecasting Model

The forecast of terminal return of Chinook Salmon to West Coast Vancouver Island (WCVI) that serves as an input to the PSC model has been done essentially the same way since it was initiated in 1999. The Panel did not have adequate time to fully understand all components of the analysis but did obtain a general understanding of the process. The forecast for WCVI relies on estimates of the return to the terminal area by age to the Robertson Creek Hatchery (RCH) obtained from both thermally marked otoliths and coded wire tags (CWT). These estimates are expanded to estimates of total run of Chinook Salmon to RCH by age using a cohort analysis. The cohort analysis accounts for both fishing and natural mortality. Two groups of sibling forecasts are made. The first group, called Prod2 by CDFO biologists, relates the estimated terminal return (which is based on CWTs of ages already observed) to the estimated total run based on the cohort analysis for the ages still at sea. The second group, called Prod3, relates the estimated total run (not the terminal run) based on the cohort analysis of ages already observed, to the estimated total run based on cohort analysis for the ages still at sea.

The same set of sibling relationships was estimated for each of PROD2 and PROD3. The relationships used the abundance in younger age classes to estimate the abundance of the subsequent age groups for that brood year. Several alternative sibling relationships were estimated. For example, the returns of age-2 fish were used to estimate the total run of age-3 + age-4 + age-5 fish. Likewise, age-3 fish were used to estimate the total run of age-4 + age-5 fish.

The two groups of sibling forecasts, Prod2 and Prod3, were then combined to produce a forecast of total run to the Robertson Creek Hatchery. The ages 3+4+5 forecast came directly from the Prod2 forecast while the ages 4+5 and age-5 forecasts were the average of the Prod2 and Prod3 forecasts for those age groups.

The combined forecast of total run for age groups was then expanded to account for the unmarked component of the return to the Somass River system, which is closely associated with the Robertson Creek Hatchery. The expansion factors were based on the annual observed proportions of RCH and Somass stocks in the RCH/Somass area.

A deterministic model was then used to remove fish harvested outside WCVI from the forecasted total run to RCH/Somass, and a maturation rate scalar was applied to estimate the number of fish by age to arrive in the terminal RCH/Somass area. For example: a harvest rate for age-3 fish harvested outside of the terminal area was applied to the age-3 + age-4 + age-5 group to remove the age-3 harvest. Then the maturation rate scalar was applied to the remainder to estimate the number of age-3 fish returning to the terminal area. Base period (1979-1982) harvest rates and maturation rates have been used for all years forecasted (1999-2016), but as noted below, these have changed.

The RCH/Somass terminal return was then expanded to the remainder of WCVI based on the ratio of observed total production (escapements plus all catches) of WCVI to observed production of the RCH/Somass system.

In recognition of reduced harvest rates in recent years compared to the base period, an alternative forecast for WCVI was provided to the CTC modeling group for input to the PSC model beginning in 2013, in addition to the forecast from the long-standing traditional forecasting method described above. For the new forecasts, the harvest rates used in the deterministic model phase were based on the recent 3-year average, whereas the maturation rate remained the same as in the base period. It appears that the forecast using the average of the 3 most recent harvest rates was selected by the CTC as input to the PSC coast-wide model for 2014 and 2015.

7.3.2 Forecasting Performance

The forecasts of West Coast Vancouver Island Chinook Salmon provided for use in the Pacific Salmon Commission (PSC) model are biased low (mean percent error = -27% without the 2013 data point and -30% with 2013 included) and are imprecise (mean absolute percent error = 42% without 2013 and 45% with 2013) (Figure 27); (calculation done by Brian Bue using data from CTC 2015b Appendix J1 for 1999-2013 and preliminary data for 2014 and 2015 obtained from John Carlile, ADF&G). The 2013 WCVI forecast of 27,338 (included in the above statistical calculations) was controversial and was not accepted by the CTC, so the PSC model produced its own forecast of 32,180 (Diana Dobson, personal communication). Both of these forecasts for 2013 ended up being way too low compared to actual returns, 84% and 82%, respectively (pair of overlapping points in top left corner of Fig. 27). Fifteen of the seventeen forecasts examined were low, although the forecast for 2007 was close (Figure 28).

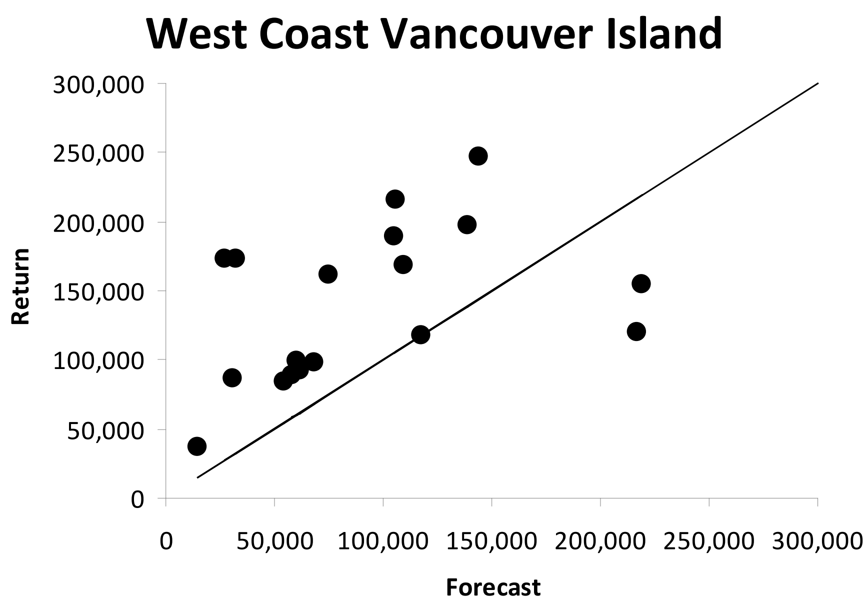


Figure 27. Comparison of the preseason terminal forecast of West Coast Vancouver Island Chinook Salmon provided by CDFO for use in the PSC model with the actual terminal return, 1999-2015. 1999-2012 data from CTC 2015b Appendix J1; 2013 (top left partially overlapping data points) from Diana Dobson CDFO and CTC 2015b Appendix J1; 2014 and 2015 are preliminary and obtained from John Carlile ADF&G).

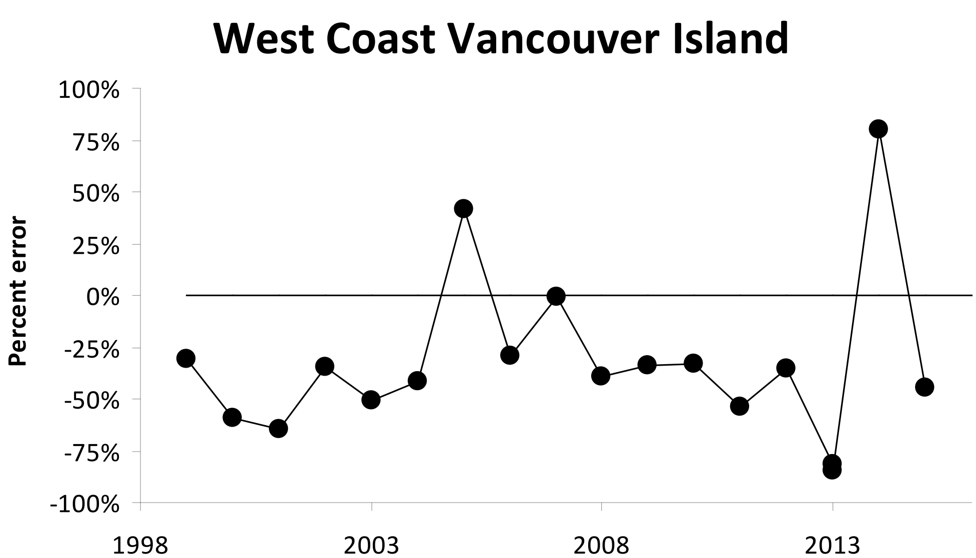


Figure 28. Percent error by year for the forecast of West Coast Vancouver Island Chinook Salmon provided by CDFO for use in the PSC model, 1999-2015. Percent error was calculated by Brian Bue as (forecast – return)/return; positive values indicate the forecast was larger than the return. The 1999-2012; data from CTC 2015b Appendix J1; 2013 partially overlapping data points from Diana Dobson CDFO and CTC 2015b Appendix J1; 2014 and 2015 are preliminary and obtained from John Carlile ADF&G).

It is clear to the Panel that the specifics of the type of forecast required for the PSC model have not been properly communicated. The lengthy use of the same forecast method, based on scalars estimated for the 1979-1982 run, contrasts dramatically with the four other stock-specific forecasts evaluated by the Panel. In all other cases, forecasters use (either implicitly or explicitly) more recent information on age composition, exploitation rates, and survival rates in their sibling relationships and cohort ratios and do not attempt to back-calculate to pre-ocean-fishery abundance. In addition, the long-term pattern of WCVI forecasts that underestimated returns should have prompted a detailed look at the forecasting methods before now. It is possible that the CDFO staff recognized the chronic underestimates but were under the assumption that the PSC model was taking them into account. CDFO staff recently attempted to address the problem by providing an alternative forecast for input to the PSC model that was based on the recent harvest history, as well as a forecast based on the traditional method, but it does not appear that any specific communication occurred between the CTC modeling group and WCVI forecasters as to how to proceed. It is imperative that both parties know what is being provided and how it is being utilized.

Recommendation 11. ***The CTC modeling group and WCVI forecasters should decide (1) which type of forecast is required from WCVI (based on base-period data or recent years, for example), and (2) the forecast performance values (bias and precision) beyond which an extensive review of forecasting methods should be triggered.***

7.3.3 Uncertainty in Parameters

Sibling Relationships

The Prod 2 sibling relationship used to forecast the age-3 + 4 + 5 grouping as well as the Prod 2 and Prod 3 relationships used to forecast the age-5 age class are problematic (Figure 29). The age-2 to age-3 +4 + 5 relationship underestimates the number of age-3 + 4 +5 fish for most years (Figure 29, pane A). This is not an uncommon problem because low-abundance age-2 fish are typically difficult to detect in a population without an intensive sampling program. Likewise the Prod 2 and Prod 3 age-4 to age-5 relationships overestimate the number of returning age-6 fish.

Recommendation 12. ***An evaluation of the WCVI sampling program should be undertaken to determine if (1) there has been a dramatic change in sample collection methods and sampling intensity over the years, and (2) whether the sample design and intensity is adequate to obtain meaningful age composition estimates. If the sample design appears to be adequate, then explore other ways to estimate the age-3 and age-6 components of the returns.***

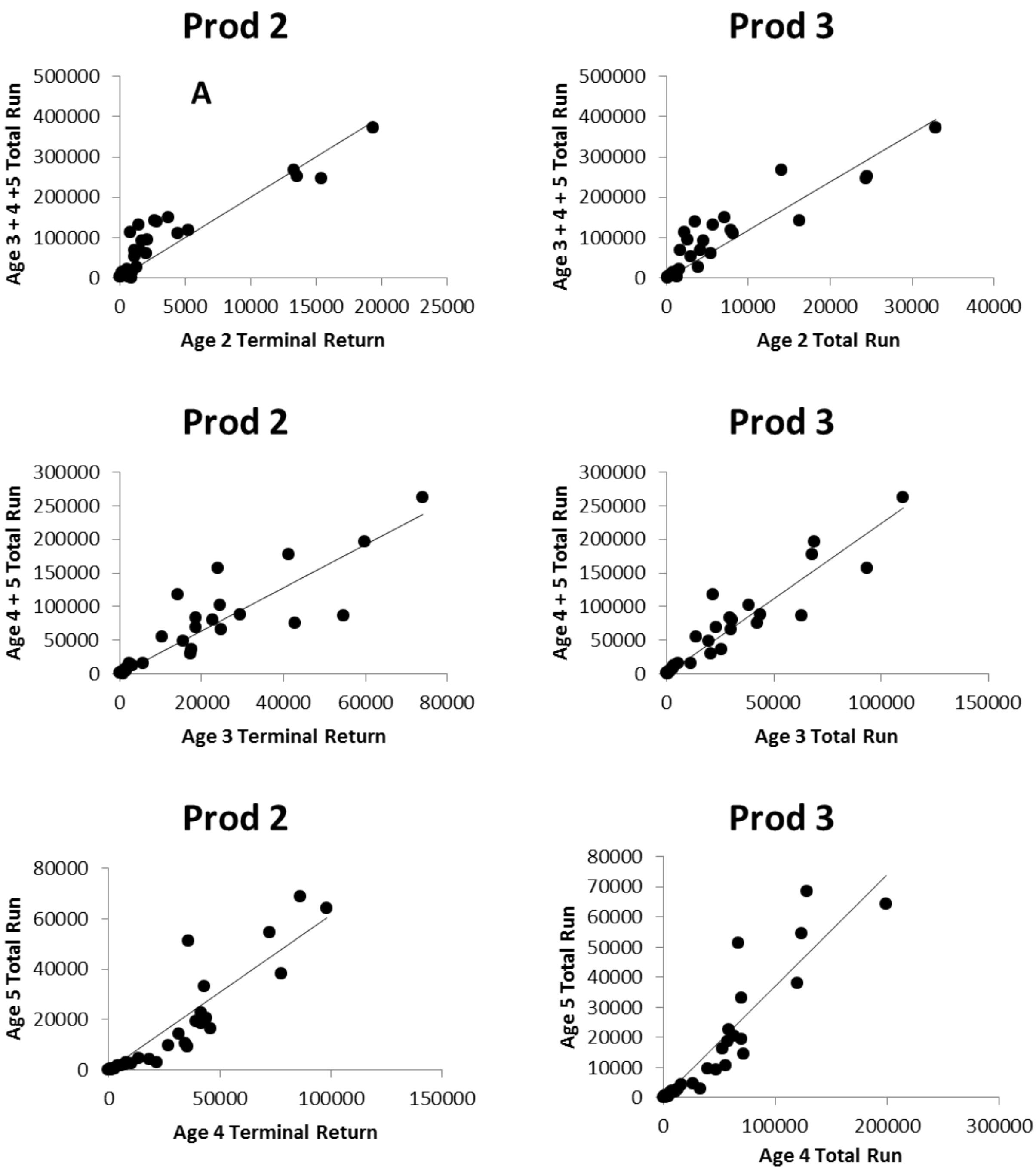


Figure 29. Sibling relationships used to forecast Chinook salmon terminal return to the Robertson Creek Hatchery, West Coast Vancouver Island. Prod 2 and Prod 3 are different groupings of the data and are explained in the text. Lines are regression lines through the origin.

<u>*Harvest and Maturation Rates*</u>

The use of base-period harvest rates (1979-1982) to estimate present day returns has the potential to negatively bias the forecast (underestimate the run). Approximately 80% of the RCH return are age-3 and age-4 fish. Two of the ages in the forecast (ages 3 and 4) are presently harvested at a lower rate in the pre-terminal areas than during the base period, and the reverse is true for age 5s (Figure 30). The use of a harvest rate that is greater than the current harvest rate would result in forecasts that underestimate total abundance in the WCVI terminal area.

The Panel did not fully examine the effect of the age-specific maturation rates on the WCVI forecast. The maturation rate is the proportion of a group of fish that will return at a given age. For example, when the base-period maturation rate for age-3 (0.17) is multiplied by the forecast number of age-3, 4, and 5 fish in the brood year return, the result is the number of age-3 fish expected to return in a given year with the remaining fish returning in future years as age-4 and 5 fish. It is highly unlikely that the maturation rate has remained constant over the past 35 years. Changes in stock productivity, oceanic thermal regimes, and food have been documented in the North Pacific and it is reasonable to expect that these changes have influenced maturation rate.

Recommendation 13. ***The use of recent harvest rates and maturation rates should be explored for the WCVI forecasting model. These analyses should estimate model sensitivity to uncertainties in these rates, and all results of these sensitivity analyses, including the associated forecasts, should be provided to CTC modelers along with estimates of uncertainty in the forecasts.***

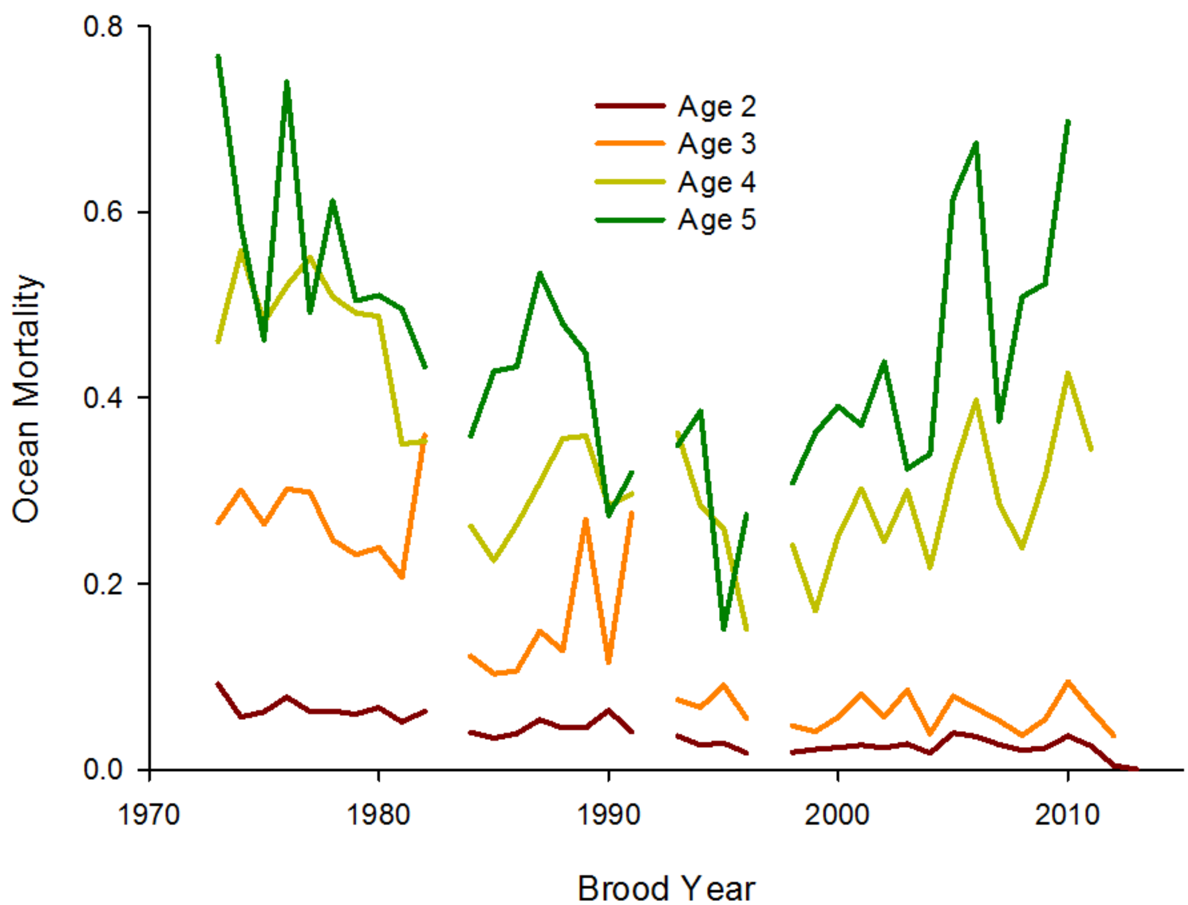


Figure 30. Age-specific pre-terminal exploitation rates of Chinook Salmon from the Robertson Creek Hatchery (Diana Dobson, CDFO, Portland workshop presentation).

Changes in Contribution of Robertson Creek Hatchery

The Robertson Creek Hatchery is the indicator stock for WCVI Chinook Salmon. One assumption underlying forecasts for WCVI is that RCH Chinook Salmon have an identical exploitation rate, maturation rate, and marine survival rate to that of other WCVI Chinook Salmon in the same years. The Panel has little information on the validity of this sweeping assumption for wild populations of Chinook Salmon, but it is clear that fish from RCH have had considerably different survival rates to age 2 than fish from other WCVI hatcheries, Conuma and Nitinat (Figure 31, right panel). Furthermore, RCH fish now contribute a much smaller portion of the total indexed WCVI terminal returns than in the late 1980s (Figure 31, left panel). Thus, the Robertson Creek Hatchery exploitation, maturation, and marine survival rates may no longer reflect those attributes for the majority of other WCVI Chinook Salmon. This issue is critically important because, as we discussed above, the RCH data are the core foundation for the WCVI forecasts. A lack of representativeness of RCH may have contributed to the poor performance of WCVI forecasts (Figure 28).

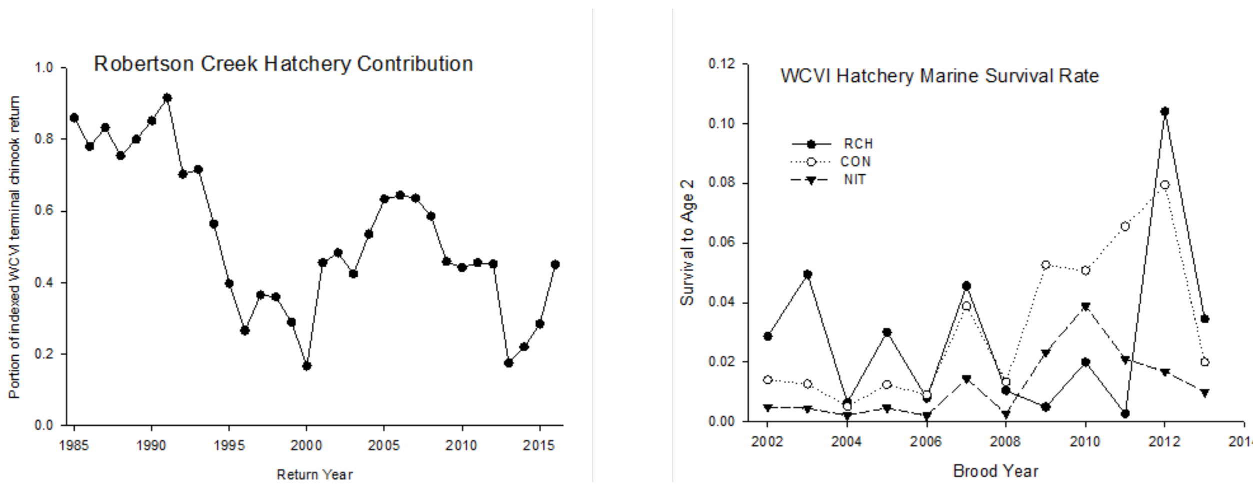


Figure 31. Contribution of the Robertson Creek Hatchery to the West Coast Vancouver Island Index of Abundance by year and the survival to age-2 for the Robertson Creek (RCH), Conuma (CON) and Nitinat (NIT) hatcheries (Diana Dobson, CDFO, Portland workshop presentation).

7.3.4 Suggestions

Explore New Models

An evaluation of the WCVI forecasting procedures should be undertaken with the goal of simplifying the process. The methods used by the Columbia River forecasters might be a good example. Presently for WCVI, coded-wire-tag, otolith, and age information are collected from the RCH stock and then expanded using a cohort analysis. Then a sibling analysis is done and abundances are further expanded for the contribution of the Somass system. Then abundances are reduced by estimated harvest and maturation rates, and again expanded for the remaining contributing stocks to WCVI. At each step of this process parameters of unknown uncertainty are applied to the basic CWT and age information to make the final forecast. In essence, terminal run information from RCH is expanded to estimate the number of fish still alive for a particular brood year and is then reduced to an estimate of future terminal return. All of these steps involve several assumptions, any of which could be substantially wrong, and it is therefore likely that a major contributor to the frequent underestimation of WCVI abundance is CDFO's assumptions about harvest rates and maturation rates. The Panel suggests that it might be easier, as well as possibly more accurate, to just forecast the terminal return for each of the three hatcheries, combine the estimates, and then expand those estimates to account for wild fish. While that endeavor would be challenging, the Panel believes it is obtainable because the forecasts for domestic management purposes appear to perform better than forecasts sent to the CTC for input to the PSC model (Diana Dobson, CDFO, personal communication).

Recommendation 14. ***Explore a different and simpler method of forecasting terminal return to WCVI. The preferred method would reduce the complexity of the forecast by reducing the number of data manipulations and number of parameters and assumptions in the forecasting procedure. As with all new methods, it should be thoroughly evaluated to determine whether an increase in performance is actually obtained in terms of bias and precision, and sensitivity analyses should be performed to determine the influence of uncertainties in model parameters.***

7.3.5 Summary

The forecasts of West Coast Vancouver Island Chinook Salmon abundance submitted for use in the PSC model have mostly been low since 1999. The Panel believes that the situation can be improved by using a new forecasting procedure that makes fewer assumptions than the present method. Considerable planning should be committed to maximize improvements in the WCVI forecasts given monetary and time constraints.

7.4 North Oregon Coast

7.4.1 Forecasting Model

The North Oregon Coast (NOC) forecast includes spawning escapements of Chinook Salmon for an aggregate of populations extending from the Siuslaw River in the south to the Nehalem River in the north. The Panelists understand that considerable improvements have been made to the program since 2008, primarily increased sampling levels for age composition, rapid turnaround in scale aging, and improvements in estimating escapement.

Sibling relationships are now used for NOC forecasts, but no data or graphs depicting the relationships were provided. The Panelists assume the relationships are generally valid, but given the changes in age composition and the scatter around sibling relationships reported above for Columbia River and WCVI Chinook Salmon, we expect that similar issues will exist for NOC stocks. Without more comprehensive data available for NOC Chinook Salmon, examination of a wider variety of forecasting models might improve forecasts (see section 6.4).

Recommendation 15. We recommend that ODFW forecasters examine log_e-log_e sibling regressions, a hybrid sibling model, and a Kalman filter estimation procedure, the latter to account for possible temporal changes in parameters of the sibling relationship.

Ethan Clemons (ODFW) illustrated the development of the 2016 NOC forecast in which biologists selected between two different forecasting methods (sibling relationships and an ARIMA time series model) for each of the seven major rivers in NOC (Figure 32). Although we encourage examination of a wide range of models, it is necessary to follow a standardized model-selection process and to document the methods and criteria used to select which model to include in the overall forecast for NOC (see section 6.4).

Recommendation 16. A list of the alternative forecasting models examined and the criteria used to select among those models for producing a forecast for the Northern Oregon Coast should be clearly stated in the forecast document provided to the PSC model group, as suggested in recommendations at the start of section 7.

Run year	Nehalem	Tillamook	Nestucca	Siletz	Yaquina	Alsea	Siuslaw	total
2016 forecast sibling ERC	12,165	16,109	9,051	7,528	15,813	45,531	40,840	147,038
2016 forecast ARIMA	9,629	12,673	8,481	5,362	6,992	27,259	28,701	99,097
							hybrid	119,374

Figure 32. Forecasts examined by ODFW staff for inclusion in the 2016 forecast of escapement to the North Oregon Coast aggregate of Chinook Salmon. Numbers highlighted in green were selected for inclusion in the 2016 forecast for the NOC aggregate stock. From the presentation by Ethan Clemons (ODFW) at the forecasting workshop in Portland, August 10-11, 2016.

7.4.2 Forecasting Performance

The forecast of North Oregon Coast Chinook Salmon provided for use in the Pacific Salmon Commission (PSC) model shows an overall tendency to overestimate abundance, and is reasonably precise (mean percent error = 8%; mean absolute percent error = 31%, but without the unusually large overestimate of the forecast in 2007, the tendency has been to underestimates, MPE = -2% and MAPE = 22%; calculations done by Brian Bue using data from CTC 2015a Appendix J1 for 1999-2013 and preliminary data for 2014 and 2015 obtained from John Carlile, ADF&G). The perceived tendency to overestimate abundance is primarily due to the forecast made in 2007 when the actual escapement was about one third of the forecast escapement (Figures 33 and 34). An improvement in forecast performance coincident with improvements in sampling, speed of scale ageing and escapement estimation in 2008 was observed beginning with the next forecast in 2009. Since 2009 there has been a tendency to underestimate the returning escapement (mean percent error = -6% for 2009-2015), but forecast precision has been greatly increased (mean absolute percent error = 14% for 2009-2015; Figure 34).

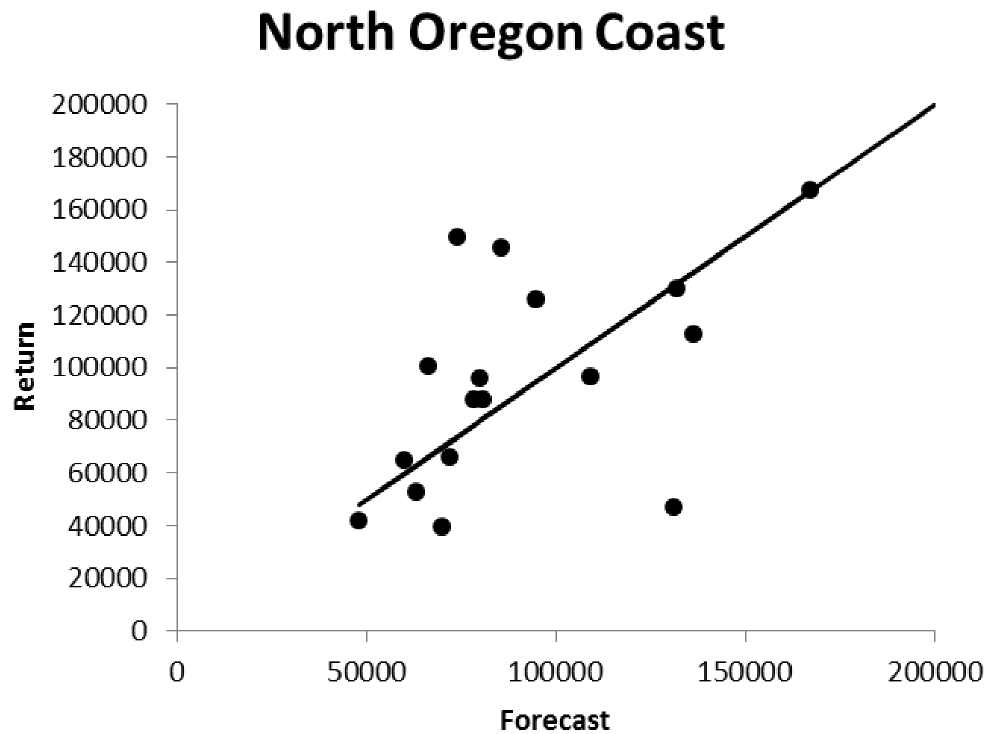


Figure 33. ***Comparison of the preseason forecast of escapement to the North Oregon Coast aggregate of stocks provided for use in the PSC model with the actual estimated escapement (return), 1999-2015. 1999-2013 data from CTC 2015b Appendix J1; 2014 and 2015 are preliminary and obtained from John Carlile ADF&G).***

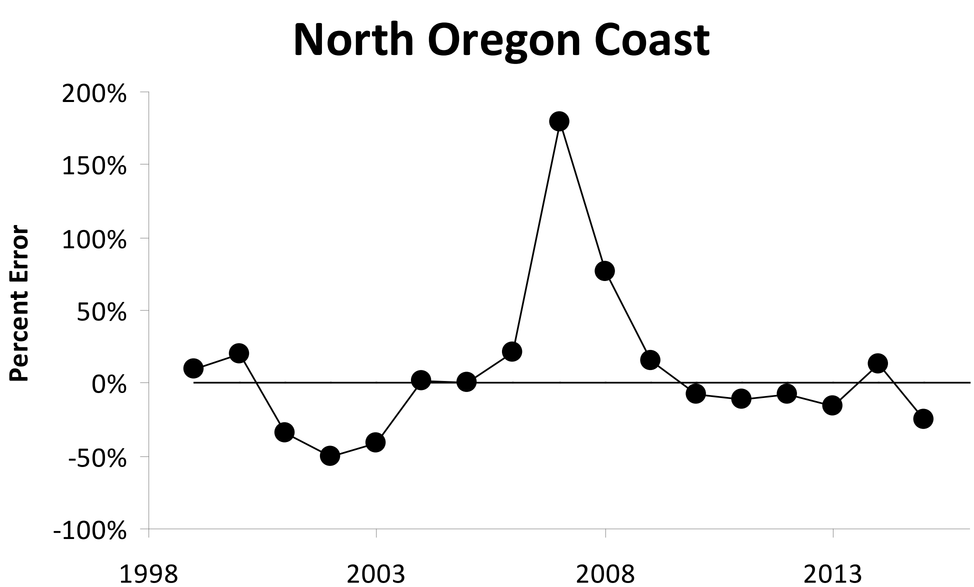


Figure 34. ***Percent error by year for the forecast of escapement to the North Oregon Coast aggregate of stocks for use in the PSC model, 1999-2015. Percent error was calculated by Brian Bue as (forecast – return)/return; positive values indicate the forecast was larger than the return. The 1999-2013 data from CTC 2015b Appendix J1; 2014 and 2015 are preliminary and obtained from John Carlile ADF&G).***

7.4.3 Uncertainty in Parameters

While there has been an increase in forecast performance since 2009, there are still many areas that could be improved. A forecast of only escapement essentially assumes that everything that affects the population in the marine environment is constant from year to year, i.e., maturation rates, exploitation rates in AABM as well as terminal fisheries, and survival from one age class to the next. It is obvious that the assumption of constant exploitation rates in the fisheries is violated because Chinook Salmon abundances vary and annual management decisions are made accordingly that can alter exploitation rates. This large source of uncertainty in forecasts for NOC can be reduced if there is concerted and increased effort made to identify stock and age composition in the various fisheries in which NOC Chinook Salmon are caught, and to do so in a timely manner such that the information can be used for forecasting (see recommendation on this topic near the start of section 7). In addition, there is an expansion made to account for escapement into areas that are not surveyed. Again it is assumed that the ratio of escapement between surveyed areas and unsurveyed areas is constant. The Panel realizes that addressing these issues will be a serious challenge if the reduction in funding occurs that we heard about at the Portland workshop.

Recommendation 17. All assumptions underlying the annual forecast, as well as data related to those assumptions, should be listed in the document provided to the PSC modelers so that everyone is aware of the forecast's strengths and weaknesses.

7.4.4 Summary

Improvements have been made for forecasting the escapement to NOC. In addition to those already mentioned, work has been under way in recent years to improve escapement estimation methods. Biologists are presently calibrating spawning ground visual surveys with mark-recapture estimates of escapement (Falcy et al. 2016), and where mark-recapture estimates are not available, visual surveys are combined with information on geomorphology and stream flow (Falcy 2015). Ethan Clemons stated at the Portland workshop that funding for field sampling has dropped precipitously in recent years, with field sampling being minimal to non-existent in some drainage basins. In addition, it now takes more time for a set of scale samples to be aged. The Panelists strongly encourage continued funding of this work so as to not lose the improvements in forecasting that have been gained in the past eight years. In addition, as the evaluation of escapement improves, researchers should be looking at expanding the forecast for NOC to become a forecast of the terminal run, where terminal harvest is also taken into account.

Recommendation 18. Continue the increased sampling in the Northern Oregon Coast for age, rapid reading of scales for age, and improvements in escapement estimation.

Recommendation 19. As the population assessment models continue to evolve, NOC researchers should determine the sensitivity of the resulting forecasts to the uncertainty in estimated parameters in the models and quantify the uncertainty in the forecasts.

Recommendation 20. If more detailed data can be obtained from terminal fisheries for NOC, the forecast for this aggregate stock should change to a terminal run forecast instead of an escapement forecast.

The software package, ForecastR, was beta tested in the NOC area in 2016 with strong acceptance by ODFW forecasters. The formal time series models are useful alternatives to sibling models, especially if data on age composition become less available.

Recommendation 21. The Panelists encourage the continued use of ForecastR for Northern Oregon Coast Chinook Salmon.

8 PACIFIC SALMON COMMISSION'S CHINOOK MODEL FORECASTS

This section on the PSC Chinook model considers two different strategies for making better forecasts of abundance. First, we make numerous recommendations that should help meet that goal if the current PSC Chinook model continues to be the main method for generating forecasts. In this option, substantial changes may be required. The second option is to eventually replace the current PSC model with one or more other models that take a different approach to making forecasts in such highly variable, complex, multi-stock, multi-region systems. We provide guidance for both options.

8.1 Forecast Performance of the Pacific Salmon Commission Model

Forecasts of Chinook Salmon abundance obtained from the PSC model after the Agency forecasts were incorporated were relatively unbiased when measured by mean percent error (MPE) for four of the five stocks in this review (Table 2). The forecast for the West Coast Vancouver Island stock was biased low (MPE=-17%) but not as biased as the Agency forecast (MPE=-30%; Table 2). While both methods of forecasting North Oregon Coast were relatively unbiased (MPE for PSC model with Agency forecast = -6%; MPE for Agency forecast = 8%) the range in MPE between the two forecasting methods (17%) was the greatest for the stocks reviewed (Table 2). As with the Agency forecasts, the forecasts obtained from the PSC model after the Agency forecasts were incorporated for the three Columbia River Chinook Salmon stocks and the North Oregon Coastal stock are at the low end of the range of MPE values observed for the chum and sockeye salmon stocks examined by Haeseker et al. (2008; Figure 14). Forecast bias (MPE) for West Coast Vancouver Island was improved and is at the lower end of the range observed for chum and sockeye salmon (Haeseker et al. 2008; Figure 14).

The precision of the forecasts (MAPE) for the five stocks from the PSC model after the Agency forecasts were incorporated was comparable to that obtained from the Agency forecasts (Table 2) and was well within the range of MAPE for chum and sockeye salmon (Haeseker at al. 2008; Figure 15).

An examination of the time series of percent error for the forecasts indicates that errors for both the Agency forecasts and the PSC model with the Agency forecasts incorporated varied together (Figure 29). Of notable interest was the time series of percent error for West Coast Vancouver Island where the Agency forecast was biased lower than the PSC model for all years between 1999 and 2012, with the opposite being true for 2014 and 2015 (Figure 29). The pattern of differences in percent error through time, closely follow the changes in methodology for developing the Agency forecast for WCVI. Namely the use of based period (1979-1982) harvest rates for the 1999-2012 forecasts followed by the use of the recent 3-year average harvest rate for the 2014 and 2015 forecasts (see Section 7.3.1). A similar but opposite pattern is present for the time series of percent error for North Oregon Coast where the forecasts obtained from the PSC model were biased lower from 2005 to 2015 (Figure 29).

The Agency forecasts and the forecasts obtained from the PSC model with the Agency forecasts incorporated were correlated, with North Oregon Coast demonstrating the most variation between the forecasting methods (Figure 30). The tendency for the forecasts obtained from the PSC model with the Agency forecast incorporated to be greater than the Agency forecasts for WCVI is evident, as is the opposite tendency for the North Oregon Coast forecasts (Figure 30).

A review of the bias and precision of forecasts obtained from the PSC model absent input from the Agency forecasts for the five stocks in this review was not performed. Conversations with John Carlile (ADF&G) indicated that the model would need to be rerun with the Agency forecasts removed in order to determine how the PSC model would forecast absent Agency input. Given the large number of possible ways the model could be examined for the five stocks (one stock removed at a time, all stocks removed, or some combination), extremely limited staff time to do the model runs, and the scope of this review, it was determined that this evaluation would best be performed at a later date.

Table 2. Comparison of bias (Mean Percent Error) and precision (Mean Absolute Percent Error) between the PSC model after the agency forecasts were incorporated and the Agency forecasts for the five Chinook Salmon stocks reviewed.

	Mean Percent Error [1]		Mean Absolute Percent Error	
	PSC Model	Agency	PSC Model	Agency
Columbia River				
Upriver Brights	-1%	1%	25%	25%
Spring Creek	-1%	8%	28%	31%
Summers	10%	5%	22%	24%
West Coast Vancouver Is. [2]	-17%	-30%	36%	45%
North Oregon Coast	-6%	8%	29%	31%

[1] Mean Percent error was calculated as (forecast-return)/return; positive values indicate the forecast was larger than the return. The 1999-2013 data are from CTC 2015b Appendix J; data from 2014 and 2015 are preliminary and obtained from John Carlile ADF&G.

[2] The evaluation of bias and precision for the PSC model from WCVI did not include information for 2013. The Agency forecast was not provided to the PSC modelers prior to the final model calibration and the effect of the missing Agency forecast on the PSC model forecast was unknown.

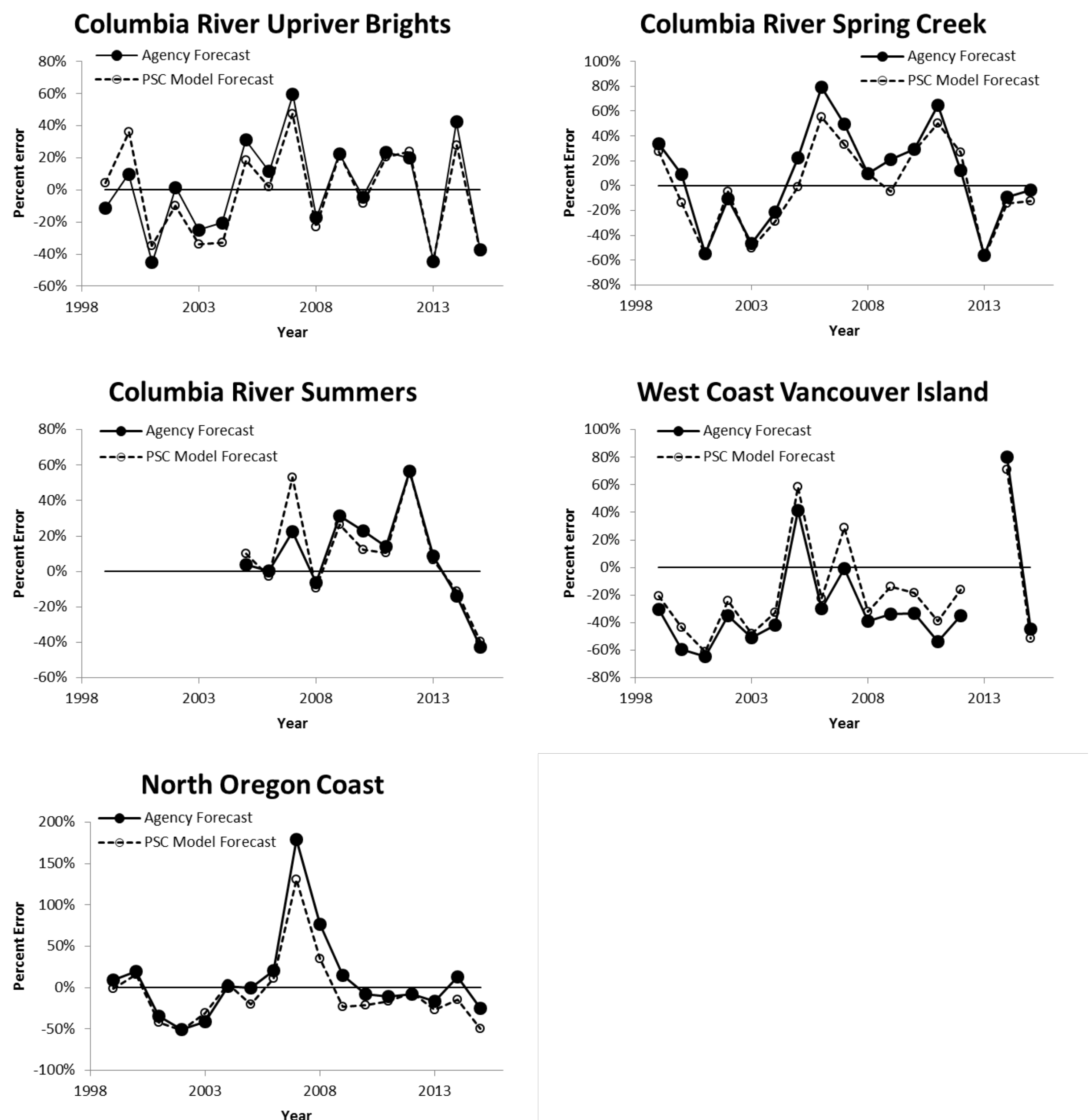


Figure 29. ***Percent error by year for the Agency forecasts and the PSC model after Agency forecasts were incorporated into the model. Percent error was calculated as (forecast – return)/return; positive values indicate the forecast was larger than the return. The 1999-2013 data are from CTC 2015b Appendix J; 2014 and 2015 are preliminary and obtained from John Carlile, ADF&G.***

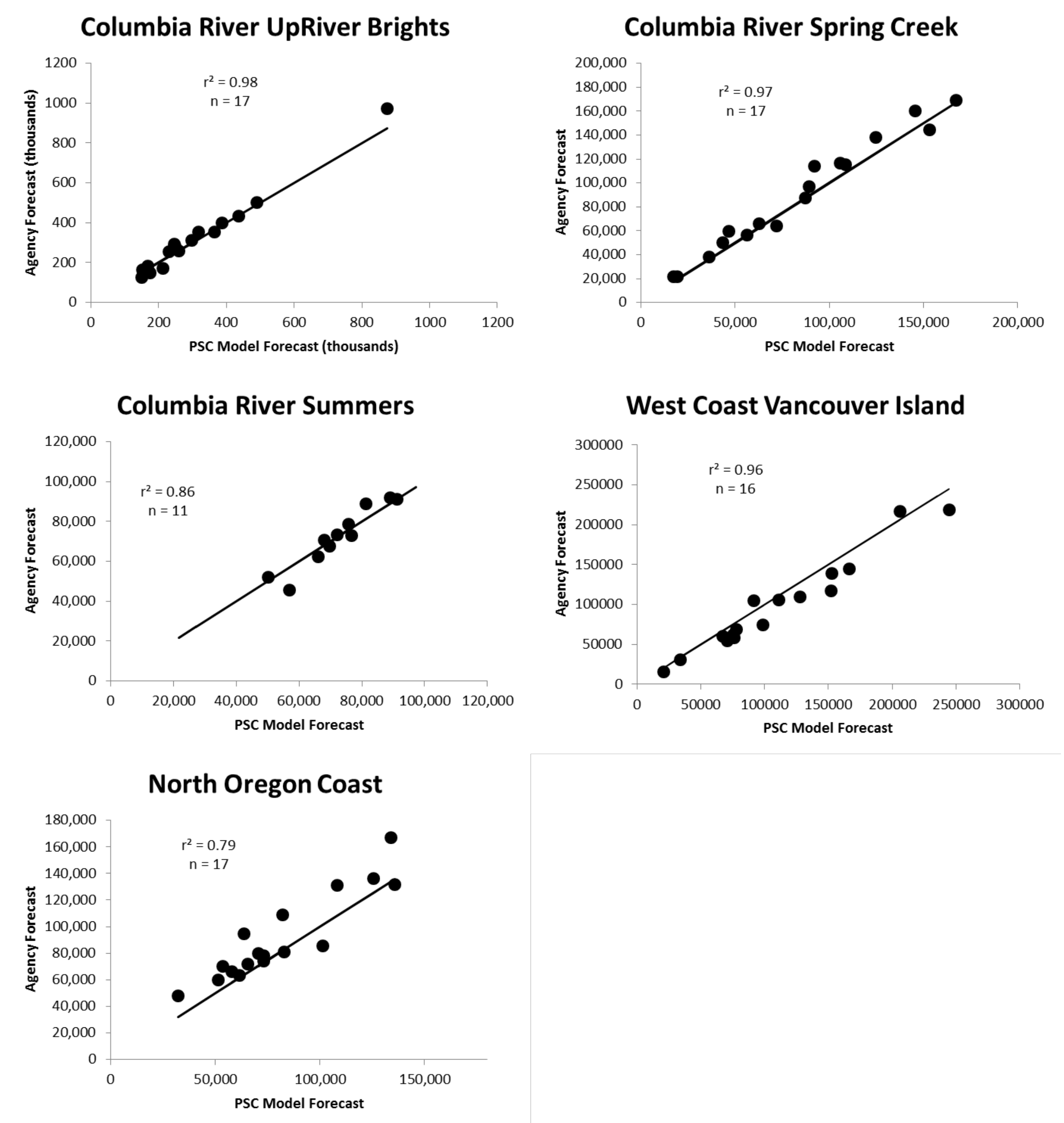


Figure 30. Comparison of the forecast obtained from the PSC model After agency forecasts were incorporated into the model with the Agency forecast. The solid line is where the Agency and PSC forecasts are equal (1:1 line). The 1999- 2013 data are from CTC 2015b Appendix J; 2014 and 2015 are preliminary and obtained from John Carlile, ADF&G.

8.2 Unclear management objectives and the PSC Chinook model

The Panel noted that the issue of uncertainty about management objectives for AABM fisheries is relevant to the PSC Chinook forecasting model in the following ways.

Example 1: As of 2016 regional biologists produced forecasts of abundance for the coming year for 23 of the 28 Chinook stocks in the PSC model. Those forecasts are sent annually to the CTC and are used as part of the input to the PSC model for its calibration and projection steps. However, the CTC does not send requests for forecasts to those regions and therefore does not provide specifications for the input data. We have already described the resulting problems in section 6.1 and have identified the need for detailed specifications from the CTC for exactly what it wants as numerical inputs from regional agencies.

Here we focus on a different issue, but it is also related to unclear management objectives for AABM fisheries. Regional agency forecasts from the 23 stocks have traditionally been produced by ranking a variety of models, including sibling age-class relationships, average abundance, and ratios of past returns to successive ages within the same cohort, with all of these estimated from data over various periods (e.g., last 3 years, last 5, etc.). The criteria such as MRE, MAPE, and r^2 that were used by agencies for ranking these models have differed among stocks, years, and even age groups within a stock. As well, in explorations of different periods of input data for estimating maturation rates to use in the PSC model, the CTC found that the best input data period often depended on the ranking criterion (CTC 2016a).

However, as noted in section 6.4, all ranking criteria are implicitly associated with different management objectives. Some criteria favor models that reduce long-term bias, some favor maximizing precision, and others try to minimize the combination of those two or even unusually large forecasting errors. Although this variety of ranking criteria is legitimate for each region's management objectives, if there are differences in ranking criteria among regions, that means that forecasts that are sent to the PSC model have different statistical characteristics (e.g., some have minimize bias, some have maximized precision). This situation will lead to confusion about what the estimated abundance indices for the three AABMs actually represent. In effect, these forecasted inputs to the PSC model are in different units. To the extent that these regional inputs influence the AABM abundance forecasts produced by the PSC model, this inconsistency in regional model-ranking criteria clouds how to interpret the point forecasts from the PSC model. As a result, when Commissioners and AABM managers look at forecast abundances for AABMs, they are currently not able to know the extent to which those forecasts reflect maximum precision or minimum statistical bias, for example. In other words, statistical uncertainty about any given abundance forecast has already been confounded with uncertainty about what that abundance actually represents (thus failing the "clarity test" of Morgan and Henrion 1990, p. 50). This appears to be an important issue that may have been overlooked by scientists and managers in the PSC, Agencies, and AABM regions.

One solution to this problem would start with clearly stated management objectives for the AABMs, which would then lead to clearly identifying which subset of possible criteria (i.e., performance measures) for ranking forecasting models would be appropriate given those objectives. Agency forecasters would then only need to generate forecasts for the one or more models that are highly ranked for those particular performance measures. If more than one performance measure is required to match the management objectives, then agencies should send to PSC modelers the forecasts from the top-ranked model (or set of near-top models) for each relevant performance measure. For instance, that might only include models that are ranked based on MPE and those based on RMSE.

Recommendation 22. ***The CTC should request each regional agency to provide to PSC modelers the forecasts of abundance for the model deemed best <u>for each of the "relevant" ranking criteria</u> (such as MRE, MAE, or RMSE), where "relevant" is defined as those that fit with stated management objectives for the AABMs .***

We emphasize that for each stock, the use by agencies of a range of ranking criteria such as MRE, MAE, or RMSE (and the best forecasting models associated with each) would be solely for the purpose of producing a set of input forecasts to send to the PSC modelers for them to conduct sensitivity analyses across management objectives that are stated for the three AABM fisheries. We would expect each regional agency to continue using its own ranking criteria for making forecasts that are to be used in management of its own terminal fisheries. Managers of different stocks will likely have their own objectives and those may even differ from objectives for AABM fisheries; for low-abundance stocks management objectives will likely be different than high-abundance objectives, for example.

To further clarify our point, the recommendation above is <u>not</u> intended to mean that the PSC modelers should carry out sensitivity analyses across ranking criteria (and hence different management objectives) <u>for</u> regional agencies. No; it is the other way around. Given clearly defined and quantified objectives for AABMs, regional agencies should conduct their own sensitivity analyses to generate forecasts for the different ranking criteria that are deemed appropriate for the management objectives, and then pass those forecasts on to the CTC. The PSC modelers can then produce AIs for the AABMs for each of those relevant ranking criteria. The latter procedure will avoid the confusion that is currently present regarding what the forecasts actually represent, as noted in the paragraph before the above recommendation.

<u>Example 2:</u> After the CTC receives agency forecasts that are based on each of the model-ranking criteria deemed relevant and appropriate to the AABM management objectives, the PSC modelers can conduct sensitivity analyses with the PSC model. For instance, the first projection run of the model might use stock-specific forecasts that were made just by each region's model that had the lowest MPE (smallest bias). That run would produce AIs for the three AABMs that reflect the least-biased stock-specific forecasts. The second projection run might use the stock-specific input forecasts based on the regions' models that minimized MAE (maximized precision), and so on through the relevant alternative ranking criteria that were deemed relevant to the management objectives for the AABMs. If consideration of model uncertainty

was desirable, projections could also be made for the second- or even third-best agency models for each relevant model-ranking criterion. This iterative procedure of parametric sensitivity analysis would then produce a range of AIs for each AABM and each performance measure (and thus each management objective) so that managers could then interpret the AIs more sensibly than at present.

Recommendation 23. ***A series of projection runs should be conducted with the PSC model to produce a range of AIs for each AABM area. These AIs would reflect the different agencies' stock-specific model-ranking criteria that are deemed relevant to AABM management objectives.***

8.3 Structural uncertainty in the PSC Chinook model

<u>*Example 1:*</u> The current PSC model assumes that Chinook Salmon are in one large pool each year for the purposes of calculating catch. In other words, there is no explicit migration from SEAK fisheries, for example, to WCVI fishing areas, as one would have in a so-called "box-car" model. Instead, catches are calculated by simply multiplying the pre-fishing abundance of the cohort (specific to each age, stock, and year) by several terms: a natural survival rate, the proportion of fish that are vulnerable to fishing gear, and the exploitation rate (1979-1982 based-period rate times another factor). However, in reality, at the high rates of exploitation that occur in some fisheries, abundance could be substantially depleted during a fishing season, making fewer fish available both later in the same area and later in fisheries further south. This depletion effect essentially means that there is competition between fisheries for the available fish. As far as the Panel can tell from discussions and the documentation, those depletion/competition processes are not represented in the PSC model, which can contribute to forecasting errors.

This is just one example of where the PSC model contains linear functions in which one or more scalars are simply multiplied by some parameter or independent variable. This basic linear structure contrasts with the workings of most fish stock assessment models, which contain numerous nonlinear functions (e.g., Quinn and Deriso 1999; Walters and Martell 2004).

Recommendation 24. ***Functionality of the PSC Chinook model might be enhanced by including, where appropriate, nonlinear relationships such as those found in many other fisheries models, including the effect of fishing on reducing the fish abundance available to subsequent fisheries during a given year.***

Forecasting errors may be reduced with such changes, although it is impossible to say by how much until those changes are made and new models are tested through numerous future years. A widely used example of a stochastic fish stock assessment model with realistic functions is the stock synthesis model of Methot and Wetzel (2013), which is described later in depth.

<u>*Example 2:*</u> The PSC model assumes that the marine spatial distribution of Chinook stocks is still the same as it was in the base period of 1979-1982. The Panel has no information on this

assumption, but we would be surprised if that spatial distribution has stayed the same during El Nino or other anomalous ocean years in which distribution has changed for other salmon species. Such changes would affect the model's estimates of exploitation rates in various fisheries. The 7-year-long International Year of the Salmon program currently being developed by the North Pacific Anadromous Fish Commission may provide an opportunity to learn more about where Chinook Salmon currently rear compared with the base period.

Recommendation 25. Effects of changes in marine spatial distribution of Chinook stocks on functionality of the PSC Chinook model need to be evaluated.

Example 3: Another type of structural uncertainty deals with lack of certainty about relationships between variables such as the effect of body size on maturation rate. The 2016 PSC model included 13 stocks (out of 28) that do not have historical age-structure data for adult returns. For these stocks, we understand that the PSC model assumes that the maturity rates for different ages are the same through time as they were in the 1979-1982 base period. This assumption is likely not valid; the age composition of the 15 Chinook Salmon stocks that actually do have such data has changed from year to year, sometimes dramatically (Appendix B of CTC 2016a), and several stocks are maturing at younger ages than in the past. Such variation is not surprising; it has been observed in other salmon species as well, including sockeye salmon, in which a general increase (not decrease) in the mean age-at-maturity across years has occurred (Pyper et al. 1999).

We recognize that the CTC has investigated alternative ways to estimate maturation rates for stocks with age-structure data, for instance, using the last 3, 5, or 9 years of data to reflect recent changes (CTC 2016a). In that 2016 analysis, they found that the 9-year average was best, combined with the previous year's EV scalar. Nevertheless, the Panel suspects that the PSC model's current lack of a link between body size and maturation rate could be one of the more influential structural uncertainties affecting the model's forecasting errors. Even for the well-documented Columbia River upriver brights, the PSC model tends to overestimate the abundance of the age-3 terminal run and underestimate age 5s, whereas for the Fraser River late stock, the PSC model produces substantial forecasting errors in both directions for escapement of age 3s, but age 5s escapements tend to be over-estimated (Antonio Velez-Espino's graphs sent to the Panel on 12 Aug. 2016).

It is well known that body size influences age at maturity in Pacific salmon. All else being unchanged, faster growth is associated with earlier maturation, as reflected by a higher-than-normal proportion of fish maturing at age 3 instead of age 4, for example. The PSC model already makes assumptions about body-size distributions in order to calculate PNV, the proportion of fish not vulnerable to fishing gear, so there should in principle be no difficulty with also using such body size distributions to calculate annual changes in maturity schedules. We learned that there are two key concerns about the current body-size assumptions as well as our proposed new approach to estimating PNV: (1) the body-size data currently influencing PNV values were gathered years ago and have not been updated, and (2) those data are fishery-specific as opposed to stock-specific, and in some regions, no body-size data have been collected for many years.

Recommendation 26. ***Sensitivity analyses with the PSC model should be used to explore different assumptions about (1) age structure for stocks without historical age composition data, (2) body-size structure used in the current method for estimating PNV, and (3) alternative structural formulations of the PSC model to calculate changes in age at maturity as a function of changes in body-size distributions. Some of those analyses could also assume various correlations with age-at-maturity schedules of other stocks.***

The range of output results will then reflect the effects of uncertainty in those structural forms of the PSC model's equations. As introduced in section 6.8, such sensitivity analyses are a routine part of many marine fish stock assessments in the U.S.A. and Canada. Such analyses are either based on a series of separate model runs, each with a different set of hypothesized forms of equations, or through inclusion of probability distributions on particular parameters that alter shapes of equations.

The Panel realizes that given the current structure of the PSC model and its calibration process, such time-consuming sensitivity analyses cannot be done in the few weeks normally available between when final input data come in and the March 1st deadline for producing forecasts for the AABMs. However, it might be possible to conduct such sensitivity analyses at other, less busy times of year. Without results from such sensitivity analyses, the PSC model's current point-estimate forecasts of abundance in each AABM do not reflect their true uncertainty, which thereby makes for less-well-informed decision making. The model's forecasts of abundance should reflect the scientist's own uncertainties about the model's structure.

Example 6: In the calibration stage, the PSC model estimates a time series of Environmental Variable (EV) values for each stock. These EV values are used to scale up or down the abundances calculated from stock-recruitment relations in the model. As we understand it, the model does not assume that EV values of different stocks can be positively correlated with one another. However, it is known for pink, chum, and sockeye salmon (Pyper et al. 2005), as well as Chinook Salmon (Riddell et al. 2013), that many separate stocks have positively correlated yearly variation in productivity (adults produced per spawner), especially stocks that have early life stages in nearby marine areas. In essence, that means that one or both parameters of the stock-recruitment relations are positively correlated across stocks and/or the residual variation around those relations are correlated. Therefore, by not using information on these shared patterns of variation, the calibration stage for the PSC model is missing an opportunity to refine parameter estimates and possibly reduce errors in forecasts of abundance in AABMs.

Recommendation 27. ***The differences between pre-season and post-season abundance indices in each of the three AABMs might be reduced by including in the PSC model***

tendencies for multiple stocks to have positively correlated time series in productivities.

This can be accomplished during the calibration process by estimating parameters of a hierarchical model that includes parameters for broadly shared parameters as well as stock-specific parameters (e.g., Mueter et al. 2002; Banerjee et al. 2004). This addition of a hierarchical framework will be challenging in the PSC model, but it has been done in other more complex salmon models (e.g., Dorner et al. 2009).

<u>*Example 7:*</u> Some components of the PSC model (such as EV values, RT scalars) probably vary in the real-world system as functions of other factors. Such processes are not currently *explicitly* accounted for in the PSC model (as opposed to *implicitly* in some catch-all adjustment parameter like EV that is estimated during calibration), yet they can change over time, not only owing to inevitable between-year variation, but also to longer term trends that may be caused by climatic change. Such changes that are not explicitly accounted for in the PSC model may partially explain the recently observed increased magnitude of both positive and negative forecasting errors. *[Incidentally, the EV term, called "an environmental factor" in some CTC documents, actually includes the net effect of all the other sources of variation that are not explicitly and dynamically included in the model-fitting process, such as shifts in age-at-maturity schedule, survival rates, and assumed harvest rates.]*

Recommendation 28. The PSC model might be improved if factors such as EV and RT were calculated as functions of other variables.

There will be additional assumptions and data requirements, but CTC modelers should at least explore how the PSC model's forecasts might be improved by explicitly including such dynamic processes. We recognize that the CTC has explored various ways to improve the PSC model (e.g., p. 130 of CTC 2014), but we suggest going further in that endeavor.

<u>*Example 8:*</u> As already noted in section 6.8, the PSC model is deterministic, that is, it does not explicitly include any sources of random variation or uncertainty. Thus, for a given set of inputs from its 11 input data files, each time the model is run, it will produce exactly the same single-value point forecasts of the abundance index in each of the three AABM areas. Not only do such point estimates fail to reflect the real-world's variability and scientists' uncertainties about the natural system, they also do not indicate to decision makers and members of the fishing industry how low or high abundances could possibly be. This complete absence of uncertainty about forecasts is unacceptable in 2016. Instead, explicit statements of uncertainty in forecasts of abundance are the norm in present-day stock assessments for other fish species.

Recommendation 29. Uncertainty in estimates from the PSC Chinook model should be explicitly represented either by making the model stochastic or running it across numerous sets of assumptions using sensitivity analyses.

For instance, CTC modelers could start by incorporating into the PSC model at least two important variance terms relevant to Chinook Salmon forecasting, maturation rate and survival rate/productivity. These could be based on empirical analyses of historical changes in those variables. If links to other variables are identified, they could be added as functions, as suggested above, but with some random error term.

Alternatively, an entirely different model, the stock synthesis model of Methot and Wetzel (2013) described below, is just one of many stochastic models that could implement this approach (also see Walters and Martell 2004).

Example 9: We expect that PSC Commissioners and area fisheries managers implicitly assume that the PSC model (like almost all fisheries models) should be able to take into account new conditions and processes. Examples would be spatial and temporal changes in Chinook productivity owing to climatic change, changes in efficiency of fishing gear and its spatial distribution since the base period of 1979-82, and changes in stock composition of fish in a given AABM. We understand that the PSC model attempts to account for such changes by assuming certain time-dependent scalars like FP (fishery policies) and RT ("a scalar to adjust the legal catch to match the observed catch under ceiling management", from John Carlile's presentation at the Portland workshop). The stage 1 calibration also fits a single adjustment factor, the time series of EV (for "environmental factors"). However, this approach of identifying various linearly related scalars is indirect and is less likely to deal adequately with new situations than a model that calculates such scalars as a function of new states (e.g., body size distributions, ocean conditions). The current PSC model also does not reflect uncertainties that are inherent in components of the equations. The larger forecasting errors produced by the PSC model in the last 5 years or so may be a result of changes in underlying dynamics that are not accounted for by the model's current functions and scalars.

In summary to this point, in the Panel's view, the PSC Chinook model's lack of representation of uncertainties in its inputs, assumptions, and outputs does not reflect the current state of the art that one might expect for a model that strongly influences management decisions in multi-million-dollar annual fisheries, especially given that they are also the subject of a major treaty between the U.S.A. and Canada.

Recommendation 30. Ideally, the existing PSC Chinook model and/or its procedures should either be tested and refined or an entirely new model (or models) should be developed.

We caution though, that the intent should not be to develop one single new model. Instead, to be consistent with the idea that there will always be structural uncertainty, several alternative structural versions of a Chinook model should be developed that differ in their assumptions. They may even differ in the amount and type of data they require. Forecasts generated with each of those models will produce a range of forecasts to illustrate the uncertainty in those outcomes. If those models are stochastic, there will be a further refinement of the probability distribution of forecasts of abundance.

One of the basic tenets of modeling is that the structure of a model should be closely tailored to its use, i.e., what managers need to know to meet their management objectives. Therefore, the following four examples of alternatives to the current PSC model should be evaluated in part by their suitability for the specific needs of the PSC. Each model can be formulated in several alternative ways using different assumptions.

The first model is the age-structured assessment model developed and tested for Copper River Chinook Salmon in Alaska by Savereide and Quinn (2004). Their "salmon catch-at-age analysis" (SCAA) model builds upon an early method of Kope (1987) and draws upon modern stock assessment methods. It uses catch-at-age data, escapements, and spawner-recruit relationships, and it is adaptable to using different types of data on selectivity of fishing, yet does not require fishing-effort data. Its parameter-estimation procedure takes into account observation (measurement) error, and various other uncertainties when making forecasts. The Panel believes that this Savereide and Quinn (2004) model may be a viable, up-to-date alternative to the current PSC model that will still be able to produce abundance forecasts for the AABMs.

The second possible alternative to the current PSC model is the "stock synthesis" model (Methot and Wetzel 2013). Stock synthesis has been used for a wide variety of species covering 61 stocks worldwide. Key features of this comprehensive modeling framework include the following. It can adapt to various amounts and types of data (including CWT information) and can estimate fishing mortality rates in the absence of a time series fishing effort data. It will forecast future abundances based in part on stock-recruitment relations. It handles multiple spatial areas and growth types, allows for changing parameters in response to environmental factors, and explicitly takes uncertainties into account to produce a range of outputs. In short, this Methot and Wetzel (2013) stock synthesis model appears ideally suited for PSC Chinook Salmon. We therefore strongly recommend that the CTC seriously consider using it, or at least comparing its suitability to the Savereide and Quinn (2004) model.

Two other types of Chinook models were developed by Morishima and Chen (2005) and Sharma (2009), the latter a statistical catch-at-age model originally proposed by Sharma and Yuen (2004) in a grant proposal that was not funded. The CTC is already familiar with these alternative models, so we will not describe them here. We were given four reasons at the Portland meeting why there has been no follow-up by the CTC with one or both of these models:

1. We were told that they would not be able to produce the abundance indices (forecast divided by abundance in the base period, 1979-1982) in column one of Table 1 in Chapter 3 of the Treaty. *We do not understand why not. It appears to us that both alternative models could produce such abundance estimates, so the CTC should take another look at those models.*
2. The Morishima and Chen (2005) model requires data on fishing effort, which "aren't readily available for some fisheries" according to a presentation at the Portland meeting. *It is not clear to the Panel whether a large or small portion of the catch is*

taken in the fisheries with missing effort information. If it is small, then the Morishima and Chen model may still be feasible.

3. It takes a huge amount of work by many people to produce the annual forecasts from the existing PSC model, and there is no time to develop other models. *Although we acknowledged this problem above in section 6, if other reporting requirements in the Treaty were reduced, or if plans to expand the PSC model to almost double the number of fisheries were postponed, CTC members would have more time to investigate alternative models. At the Portland workshop we heard doubts about the benefits of adding more stocks and fisheries to the PSC model instead of updating it in other ways.*
4. When evaluated solely on the basis of the deviation between pre- and post-season abundance indices in the three AABM areas, the PSC model was doing reasonably well up until about 2006 or even 2011 (Figure 1), and we were told that there was no perceived need to change the model. *The Panel concluded that the continuation of large forecasting errors, particularly over the last 5 years should be a strong incentive to drastically revise the PSC model to explicitly take into account the dynamics of variables such as productivity, body size, and age-at-maturity that reflect well-documented changes that are occurring in the ocean and fresh water.*

In the long term, the PSC should consider developing a formal, quantitative Management Strategy Evaluation (MSE) framework (Sainsbury et al. 2000). Such a framework entails identifying management objectives and evaluating a range of potential management options to achieve them using a set of alternative but plausible system models (not just "the best" single model) to represent uncertainties in the system's underlying dynamic processes. The output from such a MSE is a set of management strategies that are most robust to uncertainties related to model structure, parameter values, and outcomes of applying harvesting regulations (Sainsbury et al. 2000).

Management Strategy Evaluation is now considered the "gold standard" that most marine fisheries aspire to for fish stock assessment and management decision making. MSE has been used in over two dozen fisheries to derive robust management strategies, including Fraser River, Canada, sockeye salmon (Pestal et al. 2011), and pelagic as well as groundfish species, mostly in North America, Europe, South Africa, and Australia (Andre Punt, University of Washington, Seattle, personal communication).

8.4 Uncertainty in parameters of the PSC Chinook model

Example 1: Testing of the PSC model, which has now been used for decades, has never been done to see how well its parameter estimates (as opposed to abundance indices) reflect underlying true parameter values, *as determined by independently generated test data sets*. This is surprising and unsettling, given the frequency and usefulness of this approach in current fisheries research elsewhere. The Panel respects the work of the CTC, especially the members of the Analytical Working Group (AWG), who have to work within the constraints of a model that was originally built for other purposes. The cleverly designed calibration process is one way of addressing some of those constraints, but evaluation of the PSC model with independent data seems essential.

Recommendation 31. Testing of the PSC model (and all other contemplated models) should be a high priority when the Data Generating Model is released.

The lack of independent tests of the PSC model will apparently be remedied when a Data Generating Model (DGM) that is currently being developed becomes available. The DGM will have known (but hidden from the PSC model) parameter values, and will generate test data sets to feed into the PSC model. The latter model's estimates of parameters such as stock-specific EVs will be compared with known parameter values to determine the reliability of the PSC model.

Example 2: We find it unlikely that the quantitatively skilled CTC members of the AWG would have missed the following point, but based on our readings of PSC model documents and slides presented in Portland, the Panel believes that some parameter estimates could be confounded in the stage 1 calibration of the PSC model. By confounding, we mean that although a best-fit result is obtained during calibration, it may not give a unique solution; for instance, more than one combination of RT values and EVs might give equally good fits.

The following quote from CTC (2008) raises a related question about confounding of parameter estimates.

> *In other words, if the observed catch, escapement, terminal runs are reproduced correctly and the assumptions about harvest rates, survival rates and maturity rate are "reasonable", then the cohort estimate must be "right". After the model is "calibrated", you can forecast the harvest for the upcoming fisheries* (footnote on p. 46 in the Frequently Asked Questions section of CTC 2008).

The Panel is concerned with the three assumptions in this quote (harvest rate, survival rate, and maturity rate). These parameters are assumed known in order to estimate the EV factors, which essentially adjust the Ricker stock-recruitment parameters. What if EV was left set=1 and the calibration process instead estimated harvest rates, or survival rates, or maturity rates? We suspect that the resulting forecasts in the projection runs of the PSC model might be different.

In a closely related situation, it appears that the PSC model calculates for each stock the abundance of age 1 recruits from a Ricker spawner-recruit model and the EV multiplier (scalar) described above, which is initially assumed equal to 1. The model then assumes constant annual natural morality rates between successive ages, starting with 50% from age 1 to age 2, 40% from age 2 to 3, etc. down to 10% for age 5. Then in stage 1 calibration, a vector of annual EV values are estimated for each stock, assuming that all other parameters are known. But if different natural mortality rates were assumed, then the EV values would change, as would the forecasts coming out of the PSC model.

As well, a question came up at the Portland meeting about whether the calibration of the PSC model suffers from over-fitting. In other words, is it attempting to estimate too many parameters given the data? If so, then this creates a problem for the reliability of its forecasts.

Recommendation 32. ***Evaluations of the PSC model should include: (1) a check whether there is confounding of parameter estimates in the stage 1 calibration; (2) a series of sensitivity analyses/calibrations exploring alternative values for assumed age-specific natural mortality rates that might affect all other subsequent calculations and forecasts of abundance, and (3) consideration of whether the PSC model is being over-fit.***

<u>*Example 3:*</u> The Panel did not find documentation on how stock-specific Ricker stock-recruitment parameters were estimated for the PSC model. There are a few ways to do this, some of which correct for various biases. Regardless, the parameter values are assumed to be fixed, time-independent parameters in the PSC model, despite evidence that Pacific salmon have demonstrated substantial long-term trends in productivity in pink, chum, and sockeye salmon (Peterman et al. 2003; Malick and Cox 2016).

Recommendation 33. ***Documentation should be provided on the basis of estimates of Ricker stock-recruitment parameters, as well as uncertainty in those estimates. Also, some improvement in performance of the PSC model might be obtained if the AWG used a Kalman filter that allows for a time-varying maximum productivity parameter in a given stock's Ricker stock-recruitment model. That Kalman filter procedure will explicitly take into account observation error as well as natural variation.***

The resulting time series of Ricker 'a' parameters would essentially replace, and perhaps improve upon, the portion of the variation in the time series of the EV parameter that represents time-varying productivity. This Kalman filter method has been tested for its ability to track underlying changes in productivity (Peterman et al. 2000) and has been applied to pink, chum, and sockeye salmon (Peterman et al. 2003; Dorner et al. 2008; Malick and Cox 2016). Such a parameter estimation method may provide a more solid theoretical foundation for estimating time-varying parameters that should be used in the PSC model.

Example 4: In the section above on structural uncertainty, we already mentioned parameters such as maturation rates between successive ages in the context that they might be considered as functions of some variable. However, if not enough information is available to develop such functions, then uncertainty in such parameters should be considered in the PSC model by running successive sensitivity analyses. Specifically, separate model runs should be conducted for each *important* parameter such as maturity rate, with each fixed at either a plausible high, medium, or low value. Many *important* PSC model parameters might be appropriate for such analyses, including some in the 11 input data files. These could include hard-to-estimate parameters such as (1) "direct estimates of encounters [of fish with fishing gear] during CNR [Chinook Salmon non-retention] period[s] or indicators of fishing effort in the CNR period relative to the retention period", or (2) "incidental mortality rates by fishery for legal and sublegal fish that differ from those used in the base period due to alterations in gear, regulations, or fishery conduct" (CTC 2014).

Recommendation 34. Given the large number of input parameters, all possible combinations of low, medium, and high values for each parameter may be impossibly time consuming. However, only a subset of those combinations would be needed to produce a range of forecast abundances.

Sensitivity analyses conducted with only a subset of all possible combinations of parameter values is routinely done in scenario analysis or ensemble modeling (the latter term used by the Intergovernmental Panel on Climate Change, IPCC, when presenting ranges of projections given various uncertainties). In the case of the PSC Chinook model, one could first choose the best-case values for all parameters *with respect to their effect on forecasted abundance*, run the model to get the high end of the range of forecasts, and then repeat the process with the worst-case values for all parameters. Such parametric sensitivity analyses are quite common as a way to take uncertainty into account without knowing the proper stochastic equation for variation in parameters (Morgan and Henrion 1990).

Example 5: For the 6 out of the 28 Chinook PSC model stocks (as of 2016) for which agencies do not usually provide annual forecasts of abundance, the PSC model generates its own forecasts. Two of those 6 stocks have no historical age-structure data, and one of them has such age data but only up through 1993. According to the CTC, "Model stocks that do not have annual agency-generated forecasts represent about 2/3 of the catch in the NBC fishery, 1/3 in the SEAK fishery, but only 5% in the WCVI fishery" (CTC 2106b). The model assumes an age structure and then generates abundance forecasts through its assumptions about productivity. However, details of this procedure are not fully documented, and correspondence with CTC members failed to clarify it.

Recommendation 35. Additional evaluation and documentation are needed of the PSC model's methods for dealing with stocks for which age-composition data and/or

forecasts of terminal abundance or escapement are not available, given the large relative abundance of those stocks in some AABM areas.

Example 6: We also learned that proposals have been made recently to use the PSC model's own forecasts of terminal abundance for particular stocks (WCVI and Columbia River summers) instead of what some people perceive as unreliable stock-specific forecasts that are provided by regional analysts. As noted in Section 7, forecasts for those two stocks have mostly tended to be too low.

Recommendation 36. The Panel generally recommends use of stock-specific forecasts provided by agencies rather than forecasts derived solely from the PSC model in the absence of clear evidence of improvements in accuracy and precision across multiple years.

There are two reasons for this recommendation. First, as noted above, the assumptions are vague and undocumented for the PSC model's procedure for dealing with stocks for which agency-generated forecasts are either not available or are not used by the CTC. Second and most importantly, we have not seen any quantitative evaluation of bias and precision for that PSC-model procedure, so there is no reason to believe a priori that it will be any better or any worse than the current stock-specific forecasting methods (at least in terms of its influence on the estimates of abundances in the three AABMs).

8.5 Outcome uncertainty in the PSC Chinook model

Example 1: A comparison of post-season estimates of catch of Chinook Salmon with the maximum allowable catch specified in Table 1 of Chapter 3 of the 2009 Annex to the Treaty illustrates outcome uncertainty (Figure 35). In two of the three AABM fisheries, actual catch exceeded the maximum allowable in most years (CTC 2014). Although that maximum allowable catch was not a target *per se*, fishery managers presumably did not want it exceeded. The fact that it was exceeded reflects the effect of one or more of the sources of uncertainty listed above.

These three graphs (Figure 35) also illustrate the importance of showing uncertainties in forecasts of abundance indices (AIs) in AABM fisheries, which, to our knowledge, are not currently reported by the CTC to PSC Commissioners or AABM fishery managers. If the PSC model's forecasts of AI turn out to be too high, then the true AI would be to the left of its forecasted location on these graphs, causing catches to be even greater than the maximum allowable amount in the SEAK and WCVI fisheries. The reverse would be true for forecasts that are too low. Our main point is that two key sources of uncertainty interact in these important graphs, yet they are both being overlooked in the current PSC model and reporting process.

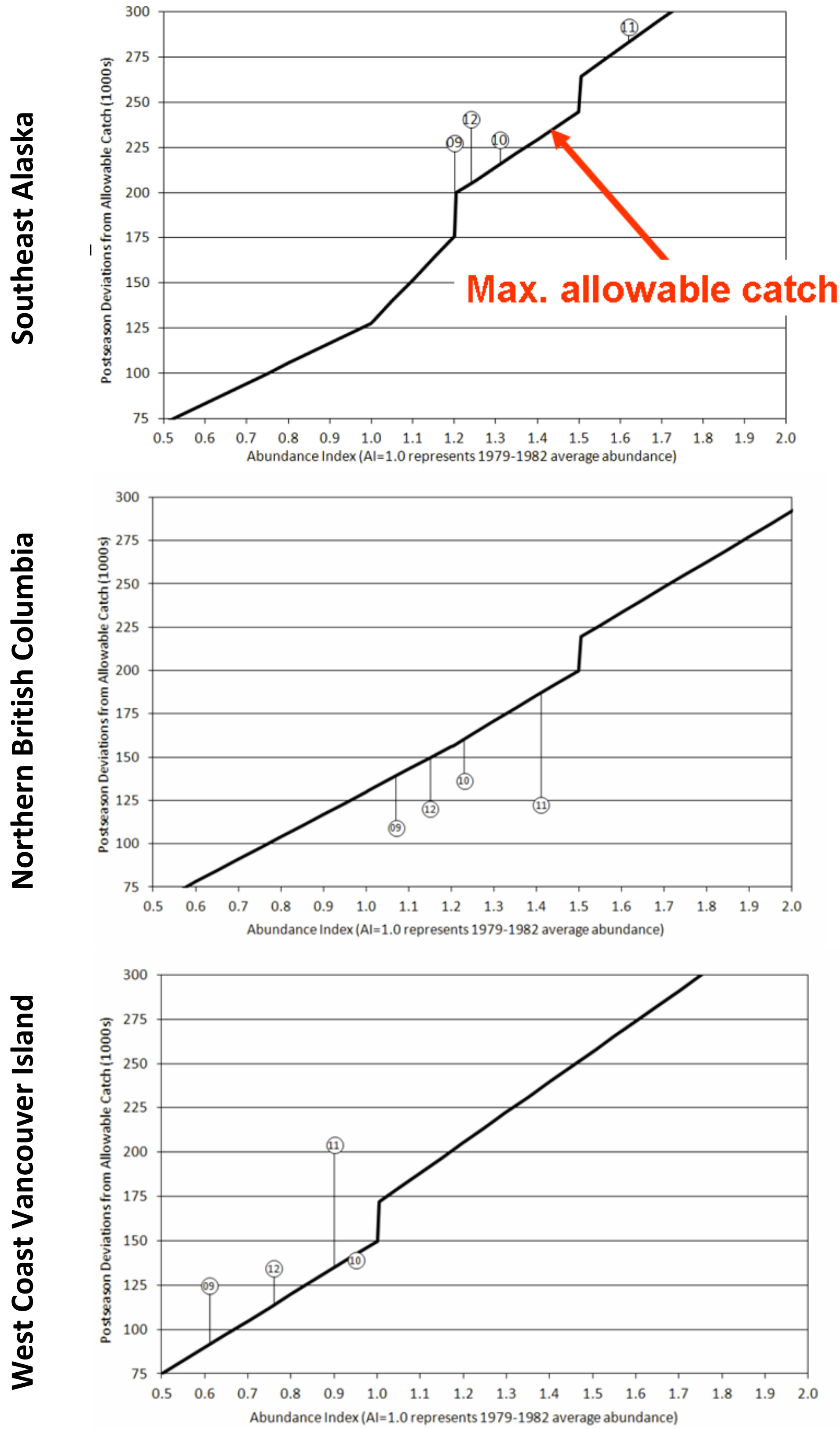


Figure 35. Post-season deviations in catch (vertical lines) from maximum allowable catch levels in the SEAK (top), NBC (middle) and WCVI (bottom) AABM fisheries for 2009-2012 (CTC 2014).

Recommendation 37 ***Considerations of outcome uncertainty (deviation between desired and realized outcomes such as catches) , as well as uncertainties in forecasts, will influence expectations of managers of these AABM fisheries when they choose annual fishing regulations.***

Failure to do so may result in not meeting management objectives, either for their own AABM areas or for the Treaty. It is beyond the mandate of this Panel to advise managers on how they should take such uncertainties into account. Suffice it to say that there is considerable experience in doing so in other fisheries, including the Fraser sockeye salmon (Fisheries and Oceans Canada 2015) and many marine fisheries around the world.

Example 2: Outcome uncertainty is relevant here because the PSC model's forecasts of abundance indices in the AABMs and terminal areas are in part based on assumptions about the exploitation rates that will occur in the forecasted year. Those forecasts should therefore reflect the difference between assumed and actual realized exploitation rates. Such differences are inevitable; managers are likely concerned about both their magnitude and direction.

Recommendation 38. ***The PSC Chinook model should take into account outcome uncertainty when making forecasts and presenting uncertainties in them.***

The Panel is unclear how best to do this, but we are confident that the CTC's AWG can empirically estimate frequency distributions of the magnitude of outcome uncertainty in past years and add those uncertainties to the PSC model.

8.6 Other issues related to the PSC Chinook model's forecasts

Example 1: We respect the care and diligence that is taken with the time-consuming annual process of calibrating the PSC model with past data and making projections. Nevertheless, we learned from some CTC members that improvements are sorely needed for that calibration process. For instance, apparently there are no pre-agreed-upon quantitative standards for deciding when a particular calibration is "sufficiently good" to be deemed as "final" by the CTC. Although seven criteria for evaluating calibration results are listed in documents such as CTC (2015, pp. 92-93), we heard that there appears to be a lack of standardized procedures for applying them from one year to the next, and for determining which model assumptions or parameter values need to be changed in subsequent calibrations if initial ones are not acceptable.

Recommendation 39. ***The calibration procedure for the PSC model should be standardized and thoroughly documented to such an extent that a new member of the Analytical Working Group could repeat previous example analyses and come to the same stopping point about which calibration is deemed "final".***

We know that in situations like this one with the PSC model, expert judgment and experience are invaluable. We anticipate that a standardized and well-documented calibration procedure will also help reduce the workload of the CTC (which we repeatedly heard is onerous) by reducing the time needed for training of new members. Such a step is also good practice for succession planning for replacing CTC members as they retire.

Example 2: As noted previously, the PSC model only produces point estimates for annual abundance forecasts in AABM areas.

Recommendation 40. ***The abundance forecasts for AABMs areas produced by the PSC Chinook model should convey to managers the net effect of all of the major uncertainties described previously -- structural uncertainty, parametric uncertainty, uncertainty about management objectives, and outcome uncertainty.***

The CTC's annual forecast AIs should be produced along with measures of uncertainty in the forecasts. We recommended the same in section 7 for stock-specific forecasts. The point estimates could still be compared to abundance values in Table 1 of the Treaty, but the addition of uncertainty measures would help with decisions based on different management objectives.

9 REFERENCES

Anderson, E.C. and J.C. Garza. 2006. The power of single-nucleotide polymorphisms for large-scale parentage inference. Genetics 172:2567-2582.

Banerjee S., B.P. Carlin, and A.E. Gelfand. 2004. Hierarchical Modeling and Analysis For Spatial Data. CRC Press, Ltd., Baco Raton, Florida, 452 pp.

Beacham, T.D., I Winther, K.L. Jonsen, M. Wetklo, L. Deng, and J.R. Candy. 2008. The application of rapid microsatellite-based stock identification to management of a Chinook Salmon troll fishery off the Queen Charlotte Islands, British Columbia. North American Journal of Fisheries Management 28:849-855.

Burnham, K.P. and D.R. Anderson. 2002. Model Selection and Multi-model Inference: A Practical Information-Theoretic Approach. 2nd edition, Springer-Verlag, New York, 488 pp.

Connors, B.M., D.C. Braun, R.M. Peterman, A.B. Cooper, J.D. Reynolds, L.M. Dill, G.T. Ruggerone, and M. Krkošek. 2012. Migration links ocean-scale competition and local ocean conditions with exposure to farmed salmon to shape wild salmon dynamics. Conservation Letters 5(4):304-312.

CTC (Chinook Technical Committee). 2008. Documentation for PSC Chinook Model, revised April 2, 2008, 105 pp.

CTC (Chinook Technical Committee). 2014. 2013 exploitation rate analysis and model calibration, Volume one. Pacific Salmon Commission Joint Chinook Technical Committee Report TCCHINOOK (14)-1 Vol. 1, Vancouver, B.C., 159 pp.

CTC (Chinook Technical Committee). 2015a. 2014 exploitation rate analysis and model calibration, Volume one. Pacific Salmon Commission Joint Chinook Technical Committee Report TCCHINOOK (15)-1 Vol. 1, Vancouver, B.C., 146 pp.

CTC (Chinook Technical Committee). 2015b. 2014 exploitation rate analysis and model calibration, Volume two: Appendix Supplement. TCCHINOOK (15)-1 Vol. 2, Vancouver, B.C., 172 pp.

CTC (Chinook Technical Committee). 2015c. Annual catch & escapement report. Pacific Salmon Commission Joint Chinook Technical Committee Report TCCHINOOK (15)-1 Vol. 1, Vancouver, B.C.

CTC (Chinook Technical Committee). 2016a. Special report. Investigation of maturation-rate estimation models and their influence on PSC Chinook model abundance indices. Pacific Salmon Commission, Report TCCHINOOK (16)–1, Vol. 1, Vancouver, B.C., 90 pp.

CTC (Chinook Technical Committee). 2016b. Chapter 3 performance evaluation report. Pacific Salmon Commission, Report TCCHINOOK (16)–2, Vancouver, B.C., 114 pp.

Dorner, B., R.M. Peterman, and S.L. Haeseker. 2008. Historical trends in productivity of 120 Pacific pink, chum, and sockeye salmon stocks reconstructed by using a Kalman filter. Canadian Journal of Fisheries and Aquatic Sciences 65(9):1842–1866, plus supplement.

Dorner, B., R.M. Peterman, and Z. Su. 2009. Evaluation of performance of alternative models of Pacific salmon in the presence of climatic change and outcome uncertainty using Monte Carlo simulations. Canadian Journal of Fisheries and Aquatic Sciences 66(12):2199-2221.

Falcy, M.R. 2015. Density-dependent habitat selection of spawning Chinook Salmon: broad-scale evidence and implications. Journal of Animal Ecology 84: 545-553.

Falcy, M.R., J.L. McCormick, S.A. Miller. 2016. Proxies in practice: calibration and validation of multiple indices of animal abundance. Journal of Wildlife Management 7(1): 117-128.

FAO (Food and Agriculture Organization). 2008. Fisheries management: 2. The ecosystem approach to fisheries. 2.1 Best practices in ecosystem modeling for informing an ecosystem approach to fisheries. FAO Fisheries Technical Guidelines for Responsible Fisheries No. 4, Suppl. 2, Add. 1, FAO, Rome, 78p.

Fisheries and Oceans Canada. 2015. Pacific Region Integrated Fisheries Management Plan; Salmon, Southern B.C., June 1, 2015 to May 31, 2016, 246 pp., available at web site: http://publications.gc.ca/site/eng/9.815513/publication.html.

Gilk-Baumer, S., W.D. Templin, and L.W. Seeb. 2013. Mixed stock analysis of Chinook Salmon harvested in Southeast Alaska commercial troll fisheries, 2004-2009. Alaska Department of Fish and Game, Fishery Data Series No. 13-26, Anchorage.

Haeseker, S.L., B. Dorner, R.M. Peterman, and Z. Su. 2007. An improved sibling model for forecasting chum and sockeye salmon abundance. North American Journal of Fisheries Management 27(2):634-642.

Haeseker, S.L., R.M. Peterman, Z. Su, and C.C. Wood. 2005. Retrospective evaluation of pre-season forecasting models for pink salmon (*Oncorhynchus gorbuscha*). North American Journal of Fisheries Management 25(3):897-918.

Haeseker, S.L., R.M. Peterman, Z. Su, and C.C. Wood. 2008. Retrospective evaluation of pre-season forecasting models for sockeye and chum salmon. North American Journal of Fisheries Management 28(1):12–29 plus Appendixes 1 and 2 available online.

Hargreaves, N.B., W. Luedke, and J. Till. 2001. Application of otolith thermal marking in British Columbia, Canada. N. Pac. Anadr. Fish Comm. Technical Report 3:19-22.

Holt, C.A. and R.M. Peterman. 2006. Missing the target: uncertainties in achieving management goals in fisheries on Fraser River, British Columbia, sockeye salmon. Canadian Journal of Fisheries and Aquatic Sciences 63(12):2722-2733.

Holt, C.A., and R.M. Peterman. 2004. Long-term trends in age-specific recruitment of sockeye salmon (*Oncorhynchus nerka*) in a changing environment. Canadian Journal of Fisheries and Aquatic Sciences 61(12):2455-2470.

Joyce, T., and D. Evans. 2000. Otolith marking of pink salmon in Prince William Sound Hatcheries 1996. Exxon Valdez Oil Spill Restoration Project Final Report (Restoration Project 99188), Alaska Department of Fish and Game, Division of Commercial Fisheries, Cordova and Anchorage, Alaska.

Kope, R.G. 1987. Separable virtual population analysis of Pacific salmon with application to marked chinook salmon, *Oncorhynchus tshawytscha*, from California's Central Valley. Can. J. Fish. Aquat. Sci. 44: 1213–1220.

Larson, W.A. 2014. AFS Genetics Section Update: parentage-based (genetic) tagging in the spotlight at the Western Division Annual Meeting. Fisheries 39(2):58.

Malick, M.J. and S.P. Cox. 2016. Regional-scale declines in productivity of pink and chum salmon stocks in western North America. PLoS ONE 11(1):e0146009. doi:10.1371/ journal.pone.0146009

Methot, R.D. Jr. 1989. Synthetic estimates of historical and current biomass of northern anchovy, *Engraulis mordax*. American Fisheries Society Symposium 6:66–82.

Methot, R.D. Jr. and C.R. Wetzel. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142:86-99.

Morgan, G. and M. Henrion. 1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge Univ. Press, 332 pp.

Morishima, G.S. and D.-G. Chen. 2005. Reformulation of algorithms for estimating fishery impacts using the CTC model. Final Report for NMFS Order: AB133F04SE0669; Reference: NFFP5000-4-00013, 63 pp.

Mueter, F.J., R.M. Peterman, B.J. Pyper. 2002. Opposite effects of ocean temperature on survival rates of 120 stocks of Pacific salmon (*Oncorhynchus* spp.) in northern and southern areas. Canadian Journal of Fisheries and Aquatic Sciences 59(3):456-463 plus the erratum printed in Can. J. Fish. Aquat. Sci. 60:757.

National Research Council. 1998. Improving Fish Stock Assessments. Committee on Fish Stock Assessment Methods, Ocean Studies Board, National Academy Press, Wash. D.C. 177 pp.

Orsi, J.A., E.A. Fergusson, and M.V. Sturdevant. 2012. Recent harvest trends of pink and chum salmon in Southeast Alaska: can marine ecosystem indicators be used as predictive tools for management? Technical report 8: International workshop on explanations for the high abundance of pink and chum salmon and future trends. North Pacific Anadromous Fish Commission, Vancouver, pp. 130-134.

Pacific Salmon Commission. 2008. Recommendations for application of genetic stock identification (GSI) methods to management of ocean salmon fisheries: special report of the Genetic Stock Identification Steering Committee and the Pacific Salmon Commission's Committee on Scientific Cooperation. Pacific Salmon Comm. Tech. Rep. No. 23: 35p.

Pestal, G., A-M. Huang, A. Cass and the FRSSI Working Group. 2011. Updated methods for assessing harvest rules for Fraser River sockeye salmon (*Oncorhynchus nerka*). Research Document 2011/133, Pacific Region, Canadian Science Advisory Secretariat, Ottawa, Canada (http://www.dfo-mpo.gc.ca/csas-sccs/Publications/ResDocs-DocRech/2011/2011_133-eng.html). Accessed 11 March 2014.

Peterman, R.M. 1982. Model of salmon age structure and its use in preseason forecasting and studies of marine survival. Canadian Journal of Fisheries and Aquatic Sciences 39:1444-1452.

Peterman, R.M., B.J. Pyper, and B.W. MacGregor. 2003. Use of the Kalman filter to reconstruct historical trends in productivity of Bristol Bay sockeye salmon. Canadian Journal of Fisheries and Aquatic Sciences 60(7):809-824.

Peterman, R.M., Pyper, B.J., Grout, J.A. 2000. Comparison of parameter estimation methods for detecting climate-induced changes in productivity of Pacific salmon (Oncorhynchus spp.). Can. J. Fish. Aquat. Sci. 57:181-191.

Pyper, B.J., F.J. Mueter, and R.M. Peterman. 2005. Across-species comparisons of spatial scales of environmental effects on survival rates of Northeast Pacific salmon. Transactions of the American Fisheries Society 134(1):86-104.

Pyper, B.J., R.M. Peterman, M.F. Lapointe, and C.J. Walters. 1999. Patterns of covariation in length and age at maturity of British Columbia and Alaska sockeye salmon (*Oncorhynchus nerka*) stocks. Can. J. Fish. Aquat. Sci. 56:1046-1057.

Quinn, T.J. II, and R.B. Deriso. 1999. Quantitative Fish Dynamics. Oxford Univ. Press, New York, 542 pp.

Riddell, B., M. Bradford, R. Carmichael, D. Hankin, R. Peterman, and A. Wertheimer. 2013. Assessment of Status and Factors for Decline of Southern BC Chinook Salmon: Independent Panel's Report. Prepared with the assistance of D.R. Marmorek and A.W. Hall, ESSA Technologies Ltd., Vancouver, B.C. for Fisheries and Oceans Canada (Vancouver. BC) and Fraser River Aboriginal Fisheries Secretariat (Merritt, BC). xxix + 165 pp. + Appendices. Available from http://www.psc.org .

Sainsbury K.J., A.E. Punt, A.D.M. Smith. 2000. Design of operational management strategies for achieving fishery ecosystem objectives. ICES Journal of Marine Science 57:731–741.

Savereide, J.W. and T.J. Quinn II. 2004. An age-structured assessment model for chinook salmon (*Oncorhynchus tshawytscha*). Can. J. Fish. Aquat. Sci. 61: 974–985

Sharma, R. 2009. Survival, maturation, ocean distribution and recruitment of Pacific Northwest Chinook Salmon (*Oncorhynchus tshawytscha*) in relation to environmental factors, and implications for management. Ph.D. dissertation, University of Washington, Seattle, 306 pp.

Sharma, R. and H. Yuen. 2004. Develop a catch at age model that incorporates both CWT data and GSI data for management purposes: an alternative approach. A grant proposal submitted to the U.S. Chinook Technical Committee for funding under the budget increment associated with the U.S. Chinook letter of agreement and abundance-based management, 11 pp.

USCTC (U.S. Chinook Technical Committee). 1997. A review of stock assessment data and procedures for U. S. salmon stocks. A Report of the Pacific Salmon Commission U.S. Chinook Technical Committee USTCCHINOOK (97)-1

Vélez-Espino, L. A., C. K. Parken, E. Clemons, R. Peterson, and K. Ryding. 2016. Automating procedures for forecasting of terminal run and escapement of Chinook, Coho and Chum salmon stocks using open-source statistical software. Proposal to the Southern Boundary Restoration and Enhancement Fund of the Pacific Salmon Commission, 48 pp.

Walters, C.J. and S.J.D. Martell. 2004. Fisheries Ecology and Management. Princeton Univ. Press, Princeton, New Jersey, 399 pp.

WDFW and ODFW (Washington Department of Fish and Wildlife and Oregon Department of Fish and Wildlife). 2016. Joint staff report: stock status and fisheries for fall Chinook Salmon, Coho Salmon, Chum Salmon, summer Steelhead, and White Sturgeon. Joint Columbia River Management Staff. Vancouver WA and Clackamas OR.

Wertheimer, A. C., J. A. Orsi, and E. A. Ferguson. 2015. Forecasting pink salmon harvest in Southeast Alaska from juvenile salmon abundance and associated biophysical parameters: 2014 Returns and 2015 Forecast. NPAFC Doc. 1618. 26 pp. National Oceanic and Atmospheric Administration (NOAA), National Marine Fisheries Service (NMFS), Alaska Fisheries Science Center, Auke Bay Laboratories, Ted Stevens Marine Research Institute (Available at http://www.npafc.org).

10 APPENDICES

Appendix A – Executive Summary of Review Process

Review of Chinook abundance forecast methodology:

Summary of process and timelines

Issued by the Executive Secretary
May 24, 2016

Purpose

Various concerns have been raised about preseason forecasts provided by agencies as input to the annual calibration procedure of the Pacific Salmon Commission (PSC) coast-wide chinook model and how these are incorporated into the calibration procedure which ultimately results in the Abundance Indices (AIs) and total allowable catches for the three *Aggregate Abundance Based Management* **(**AABM) fisheries. Thus, the Commission approved a process and timeline for an independent technical review of agency pre-season abundance forecasts and their application in the PSC's Chinook Technical Committee (CTC) coast-wide model. The review will include three methods for predicting stock abundance (agency forecast, CTC model calibration from agency forecast, and CTC model forecast absent agency forecast) and:

1) Evaluate the bias and precision of each method in predicting the pre and post-season abundance (Abundance Index (AI));
2) Provide advice on the strengths and weaknesses of each method; and
3) Suggest improvements to current agency pre-season forecast methods for predicting stock abundance.

Approach

An Independent Expert Panel ("the Panel") will be established consisting of three members, one proposed by Canada and two by the U.S. Section to review agency and PSC model pre-season forecasts.

The Panel will identify appropriate criteria that should be used to evaluate accuracy and precision of pre-season agency forecasts and model projections, and conduct a detailed review of methodology and subsequent performance of agency produced and chinook model produced forecasts. The Panel will identify which years should be used to evaluate forecast performance, for example the current chapter years 2009 to 2015 or as deemed helpful by the Panel. The review should describe current methods used for agency and model forecasts, how the agency forecasts are incorporated into the chinook model, provide a diagnosis of the deviations between model and agency forecasts, and suggest improvements on current methods to minimize error and improve precision compared with the final-post season estimates.

The project will have four stages:

i. Workshop/Information Collection: The Panel will hold a workshop/meeting in person with officials directly responsible for the selected agency forecasts and the PSC's CTC Chinook model. Agency staff will provide technical/analytical information about the domestic agency forecasts and annual run reconstruction methods. At this session, a document/presentation will also be made on the PSC Chinook Model in order to familiarize the panel members with the Chinook Model calibration procedures:

   a. incorporating the agency-provided forecasts,
   b. generating forecasts for stocks for which an agency forecast is not available,
   c. calculate the pre-fishery cohort abundances which are then allocated to ocean and terminal fisheries.

ii. Develop a draft report: The panel members will review the information and develop a draft report according to the above mentioned requirements.

iii. Review process: The panel will provide the draft report to the agencies in order to check facts, identify errors, and to avoid misunderstandings regarding feasibility of the recommendations to improve the agency forecasts. The agencies will provide comments back to the Panel within two weeks. Should the Panel wish to discuss comments with agencies: phone calls, virtual meetings or ad hoc small meetings should be used to minimize costs whenever possible.

iv. Update to the Commission: A brief update on progress and elements of the draft report will be provided at the Fall Session of the Commission in Vancouver, B.C., 3-7 October 2016. The panel will provide the update virtually or in a way that minimizes costs.

v. Finalize report: The panel will finalize the report and transmit it to the Commission by November 14, 2016.

Stocks to be considered in the review

The Commission has selected the following stocks for the review to manage the Panel's work and to focus on stocks which have an impact on AABM fisheries and those in which there are performance issues:

i) Columbia River Upriver Brights
ii) Columbia River Summers
iii) Spring Creek (Columbia River)
iv) Oregon Fall Coastal
v) West Coast Vancouver Island

Timelines

i) Following adoption by the Commission, the Secretariat will alert the agencies affected to the review process and timeline and request their support in the collection of pertinent information to be provided to the Panel in advance of an information collection workshop to be held in early summer. The Panel will confirm dates with the Secretariat and designate a primary panel contact.

ii) The Panel will provide a draft report for fact-checking review to the relevant agency staff by mid-September 2016. Agencies will have two weeks to review the draft and provide comments to the panel by October 1, 2016.

iii) Should they wish, the Panel may follow up with the agencies' staff to discuss any of the errors identified in their fact-check review.

iv) The Panel will provide its final report to the Commission by November 14, 2016.

Implementation

The PSC Executive Secretary will work closely with the respective National Correspondents to implement the process by:

- Engaging the panel members identified by the Parties through contract for work.

- Arranging the necessary meetings with agency staff, including the initial in-person meeting in Portland, Oregon in early summer 2016 and the other virtual meetings.

- Managing the payments to the panel members and tracking progress.

Panel members are expected to arrange meetings amongst themselves however they may request the assistance of the Secretariat including providing webinar access.

Budget/Funding

It is proposed that the panel members may use up to 20 days to carry out the review. Should the panel members determine that more days are required to complete the work, while remaining within the timeline for reporting to the Commission, they should notify the Commission by the end of August with a detailed rationale. The Parties will be responsible for the costs of their respective staff in the process. It is envisioned that the panel's costs will be provided by the Parties: two members by the U.S. and one member by Canada.

The sources of funds will be identified internally by the Parties. As determined by each Party, the respective agencies may be offered funding or asked to bear the costs of their experts' participation in meetings including any required travel expenses. Considering that four of the five stocks subject to review are located in Oregon, the first in-person meeting could be held in Portland, Oregon. To contain costs, any further meetings should be held virtually, however if the Panel believes in person meetings are required then they should be held in a location that meets the financial needs of all agencies.

Appendix B – Independent Panel

Brian Bue, Owner, Bue Consulting LLC. Brian received his undergraduate and graduate degrees from the University of Alaska, Fairbanks. He has provided consultations on Arctic-Yukon-Kuskokwim salmon projects through the Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative. He is retired from the Alaska Department of Fish and Game/Commercial Fisheries Division, and was the project leader for forecasts for Bristol Bay sockeye salmon.

Randall M. Peterman, Professor Emeritus, Simon Fraser University. Randall held a Canada Research Chair in "Fisheries Risk Assessment and Management" from 2001 through 2012 and specialized in quantitative methods to improve fisheries management. His research focused on: (1) fish population dynamics, (2) uncertainties affecting conservation risks and management decisions, and (3) reducing uncertainties.

Ray Beamesderfer, Fish Science Solutions. Ray is a fisheries expert in the Pacific Northwest with specialties in fishery management, statistical analysis, biological assessment, life history, and effects modeling, among other services. Ray last worked for the Oregon Department of Fish and Wildlife in 2000, and since has held positions with a variety of consulting firms.

Appendix C – Portland Workshop, 10-11 August 2016

Agenda

<u>**Day 1, Wednesday the 10th of August 2016, 8:30 AM – 5:00 PM**</u>

Start 8:30 AM

1. Welcome and logistics -- ***Ray Beamesderfer, Meeting Facilitator***
   - Introductions

2. The Panel's process -- ***Panelists (Randall Peterman, Brian Bue, Ray Beamesderfer)***
   A. Terms of reference - focus on technical scientific issues
   B. Five focal Chinook stocks
   C. Time line
   D. Conceptual framework

3. Participant Perspectives **- All**

4. General overview -- ***Gayle Brown (CFDO) & John Carlile (ADFG)***
   A. How the Pacific Salmon Treaty sets the context for the Panel's work on forecasting
   B. For each major fishery, what is the percentage catch composition by stock (time series)?
   C. For each of the 5 Chinook stocks in this review, what percentage of their catch occurs in each major fishery (time series)?
   D. Very broadly, how Chinook forecasts are fed into the annual regulation-setting process

5. Pacific Salmon Commission Chinook model ***-- John Carlile (ADFG), Gayle Brown (CFDO), and Antonio Velez-Espino (CDFO)***
   A. Role of model in implementation of the Treaty
   B. Past forecasting performance of the Chinook model
      i. Comparisons with annual post-season estimates of abundance
      ii. Comparisons with stock-specific agency forecasts for those stocks
   C. Chinook model
      i. General structure
      ii. Key assumptions and evidence for how well they are supported by data
   D. Annual calibration procedures for the Chinook model
      i. Incorporating agency forecasts
      ii. Generating forecasts for stocks where an agency forecast is <u>not</u> available
      iii. Again, key assumptions and evidence for how well they are supported by data
   E. Estimation of the pre-fishery cohort abundances, which are then allocated to ocean and terminal fisheries
   F. Estimation of post-season abundances against which forecasts have been compared
   G. Alternative versions of the Chinook model that have been explored
   H. Suggestions for how to improve forecasts
   I. Questions from Panel

6. General discussion
   A. Further suggestions for how to improve forecasts for the CTC's Chinook model
   B. Alternative performance measures, such as minimizing frequency and/or magnitude of large (>X) forecasting errors, being more concerned about over-estimates than under-estimates (or vice versa), etc.

7. Presentations for each of the five Chinook stocks, subject to re-ordering.

[Please note: The Panel respectfully requests that all presentations below (1) begin with describing the historical performance of forecasts (at a minimum, time series as well as measures of bias and precision) in order to set the context for your descriptions of (2) forecasting methods, (3) why certain years were used as inputs to forecasting methods, (4) key assumptions and limitations/weaknesses of existing forecasts and the underlying data, (5) which other forecasting methods have been explored, and (6) suggestions for how to improve forecasts.]

   A. West Coast of Vancouver Island (WCVI) (**max. 45 min.**) -- ***Diana Dobson (CDFO)***
      i. Agency's presentation
      ii. Questions from Panel

5:00 PM -- [End of first day approximately here]

Day 2, Thursday the 11th of August 2016 - 8:30 AM – 4:30 PM

8. Continue stock-specific presentations
   A. West Coast of Vancouver Island (WCVI) - continued
      i. Agency's presentation
      ii. Questions from Panel

   B. Columbia River Fishery Orientation **– Jeff Whisler (ODFW)**

   C. Columbia River Upriver Brights -- **Steve Haeseker (USFWS)**
      i. Agency's presentation
      ii. Questions from Panel

   D. Spring Creek hatchery **– Steve Haeseker (USFWS)**
      i. Agency's presentation
      ii. Questions from Panel

   E. Columbia River Summer run -- ***Stuart Ellis (CRITFC)***
      i. Agency's presentation
      ii. Questions from Panel

   F. Northern Oregon Coast fall **-- *Ethan Clemons (ODFW)***
      i. Agency's presentation
      ii. Questions from Panel

9. ForecastR package **-- *Antonio Velez-Espino (CDFO)***

10. General discussion
   A. Suggestions for how to improve forecasts for the five focal stocks
   B. Experience with combining results from multiple forecasting models
   C. Alternative performance measures
   D. How good is good enough for the forecasts?

11. Next steps and action items, including the Panel's further requests for information
   - Forms for participants

12. Wrap-up – ***Panelists***

Attendance

Name	Agency
Alan Byrne	IDFG
Antonio Velez	CDFO
Ben Cox	WDFW
Bob Clark	ADFG
Brian Bue	Panel
Christine Mallette	ODFW
Diana Dobson	CDFO
Ethan Clemons	ODFW
Gayle Brown	CDFO
Jeff Whisler	ODFW
John Carlile	ADFG
John North	ODFW
Jon Hess	CRITFC
Lisa Harlan	WDFW
Marianne McClure	CRITFC
Matt Falcy	ODFW
Randall Peterman	Panel
Ray Beamesderfer	Panel
Robert Kope	NMFS
Robin Ehlke	WDFW
Roger Dick, Jr.	YN Fisheries
Ron Roler	WDFW
Steve Haeseker	USFWS
Stuart Ellis	CRITFC
Tim Dalton	ODFW
Tommy Garrison	CRITFC

The Panel's Process – Randall Peterman (Panel)

The Panel's Process

Pacific Salmon Commission's "Workshop on Chinook Forecasting",
Portland, Oregon, 10-11 August 2016

1

A. Terms of Reference for the Panel

- Conduct a detailed technical review of methods and subsequent performance of:
  - -- Agency-produced forecasts
  - -- Forecasts from the CTC's Chinook model (with and without agency forecasts as inputs)
- Evaluate bias and precision of each method based on pre- and post-season abundances

2

Terms of Reference (continued)

- Identify appropriate criteria that should be used to evaluate bias and precision
- Provide advice on strengths and weaknesses of each forecasting method
- Suggest improvements on current methods

3

4

Common objective:
Produce the best possible forecasts.

Proactive workshop: collaborate to seek improvements

5

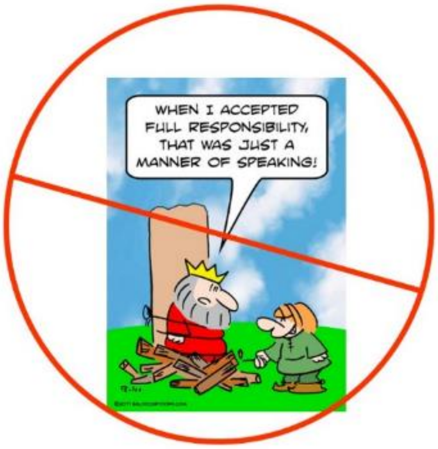
Common objective:
Produce the best possible forecasts.

Proactive workshop: collaborate to seek improvements

Patience

Questions

6



7

B. Five Chinook stocks

- West Coast Vancouver Island
- Columbia River Upriver Brights
- Columbia River Spring Creek Hatchery
- Columbia River Summers
- Northern Oregon Coastal fall Chinook

8

C. Time line for Panel's work

- Started **15 July 2016**
- Meeting in Portland **10-11 August**
- Draft report by **15 September**
  - Will be sent to PSC and area agencies to check facts, identify errors, etc.
  - By **30 September**, CTC and area agencies to provide comments to Panel on draft report
- Verbal report to PSC **3-7 October**
- Final report by **14 November 2016**

9

D. Conceptual Framework for Panel's Review: Sources of Uncertainty

Pre-season forecasts of abundance

Post-season estimates of abundance

Forecasting errors result from incomplete understanding of real world

10

D. Conceptual Framework for Panel's Review: Sources of Uncertainty

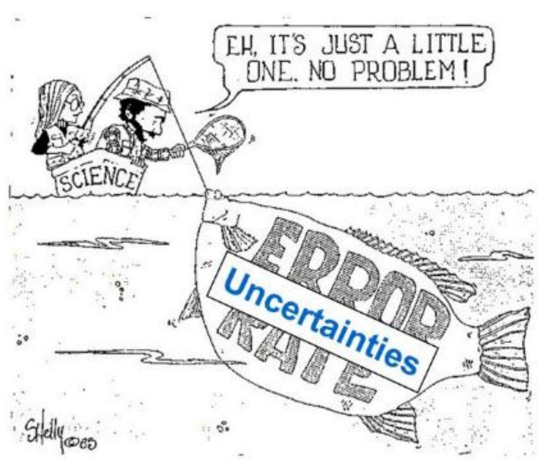
Pre-season forecasts of abundance

Post-season estimates of abundance

Forecasting errors result from incomplete understanding of real world

Hypotheses (models): regional forecasting models and PSC Chinook model (all models include uncertainties)

11



12

Uncertainties in hypothesis/assumptions

- Structural uncertainty
- Parametric uncertainty
- Unclear management objectives
- Outcome uncertainty

13

Uncertainties in hypothesis/assumptions

- Structural uncertainty -- forms of equations

14

Uncertainties in hypothesis/assumptions

- Structural uncertainty

"Now that desk looks better. Everything's squared away, yes sir, squaaaaaared away."

Uncertainties in hypothesis/assumptions

- Structural uncertainty

- Linear or nonlinear ?
- How include effects of body size, ocean conditions, size-selective fishing, ...?
- ...

16

Uncertainties in hypothesis/assumptions

- Structural uncertainty
- Parametric uncertainty

- Imperfect data from observation error
- Variations over time and/or space in:
  - -- Maturation rate
  - -- Survival rate / productivity
  - -- Efficiency of fishing gear
  - -- Proportion vulnerable
- Input variables (catches, CWT, forecasts, ...)

17

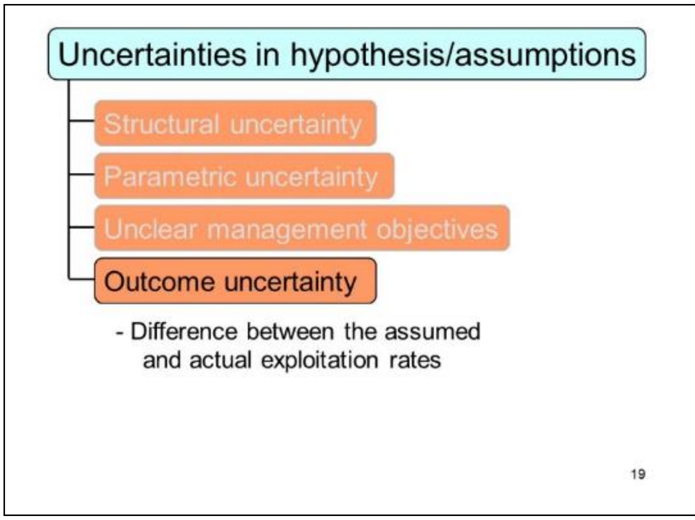
Uncertainties in hypothesis/assumptions

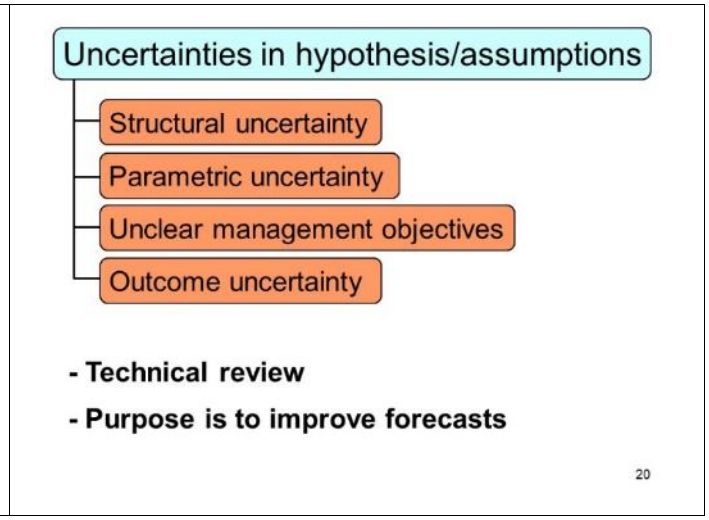
- Structural uncertainty
- Parametric uncertainty
- Unclear management objectives

Rank order of alternative forecasting models can be affected by ranking criteria:

- Minimize bias
- Maximize precision
- Minimize prob. of large errors (e.g., >30%)
- Minimize frequency of over-forecasts (or under-forecasts)

18


Uncertainties in hypothesis/assumptions
Structural uncertainty
Parametric uncertainty
Unclear management objectives
Outcome uncertainty
- Difference between the assumed and actual exploitation rates
19


Uncertainties in hypothesis/assumptions
Structural uncertainty
Parametric uncertainty
Unclear management objectives
Outcome uncertainty
- Technical review
- Purpose is to improve forecasts
20

Expectations of non-panel participants at the Portland workshop – Facilitated Discussion

Table 3. Responses to the question of "what is the single most important problem or issue for improving forecasts?"

1.	Unclear management objectives. Uncertainty needs to be described.
2.	Model is designed to minimize pre and post season AI's, not fit to the inshore population. For the two types of forecasts, agencies' forecasts of terminal runs/escapement and CTC model's forecasts of abundance indices in AABMs, a model selection process is critical and should be based on retrospective analysis.
3.	Would like to better forecast large runs that are at levels above beyond the historical data.
4.	Persistent recent bias in WCVI forecast, possibly an environmental effect. There is a difference between agency forecasts used for regional management purposes and the forecasts presented for to the CTC model. Interested in other model inputs besides forecasts. Helpful to put some thought into other inputs. Inseason updating has not been a part of Chinook management.
5.	Interested in better age composition information. Also interested in the maturation process. There is no process in place for bringing new information into the PSC model's forecasting process, e.g., winter troll fishery as an early indicator of a big run.
6.	Uncertainty in the AI's is absent but should be presented. Different ways to approach the estimation of uncertainty.
7.	Interested in what stocks are present in the fisheries. Need an appropriate forecast for the question; forecasts from different regions are incompatible because of different management objectives. Forecasting record- high returns is a concern.
8.	Difference in maturation rates between natural and hatchery stocks, yet there is heavy reliance on hatchery data. Lack of good data on age composition of escapements in some stocks; there are strong and weak assessments. Present CTC model is deterministic.
9.	Quality and lack of data: Less than half of the stocks feed age-structured forecasts of abundance into the CTC model.
10.	Appropriate models unclear -- with recent observations of abundance above the historical range, what do we assume for the shape of the function?; What are we going to do with uncertainty estimates?
11.	The five stocks selected for the this review drive fisheries, yet there is no way to judge the quality of stock-specific forecasts being handed to CTC members and SEAK AABM managers. Wants documentation of uncertainty and performance, as well as methods of the stock-specific forecasts.

Overview of the PST, the PSC Chinook Model and How Agency Forecasts are Incorporated into Model

John K. Carlile, Alaska Department of Fish and Game
Gayle Brown, Canadian Department of Fisheries and Oceans

1

Very Brief History of the Pacific Salmon Treaty

- 1985 US-Canada Pacific Salmon Treaty
- All species of Pacific salmon and fisheries from Cape Falcon, Oregon to Cape Suckling, Alaska
- Generally accepted by the late 70s-early 80s that Chinook stocks were depressed coast-wide
- Recognition that a Chinook rebuilding program was needed
- Rebuilding assessment tool was needed as well

2

Chinook Technical Committee (CTC)

- Reports to the Pacific Salmon Commission (PSC); tribal & agency representatives from AK, BC, WA and OR.
- Produces yearly reports on Chinook catches and escapements, Coded wire tag (CWT) exploitation rate analysis (ERA) and PSC Chinook Model calibrations.
- Produces other analyses and reports at the request of the PSC Commissioners.
- The Analytical Work Group (AWG) - subgroup of the CTC - produces annual ERA, Chinook Model calibration.

3

History and Development of the PSC Chinook Model (Part 1)

- Cohort Analysis model
- 1983 – Visicalc Spreadsheet with one stock
- 1984 – Apple Basic with one stock model
- 1985- MS Basic, 5 stocks, 10 fisheries
  - ➢ Evaluation of management strategies (catch ceilings, harvest rates etc.)
  - ➢ Basis of the 15 year rebuilding program (catch ceilings)

4

History and Development of the PSC Chinook Model (Part 2)

- 1988 - MS QuickBasic, 27 stocks, 25 fisheries
- 1990- 28 stocks, 1993-29 stocks, 1994-30 stocks
- 2004 - MS VisualBasic 6, 30 stocks, 25 fisheries
- 2010 - MS VB.NET, 30 stocks, 25 fisheries
- 1999- Model changed from a rebuilding assessment tool to a management tool

5

SEAK Stock Composition Example

FISHERY	SE ALASKA ALL GEAR				
	2015	Average (1985-2014)			
Model Stock	% of Fishery Catch	% of Fishery Catch	% of Stock Catch	% of Stock Total Return	Associated Escapement Indicator Stocks
Columbia Upriver Bright	20.97%	18.09%	26.61%	18.32%	Columbia Upriver Bright
North/Central BC	12.63%	16.51%	20.04%	10.53%	Yakoun Nass Skeena Area 6 Index Area 8 Index Rivers Inlet Smith Inlet
WCVI Hatchery	10.66%	15.50%	52.64%	18.12%	NA
Oregon Coastal North Migrating	7.75%	11.46%	34.01%	15.14%	Oregon Coastal
Mid-Columbia Brights	11.80%	6.56%	20.15%	14.00%	Not Represented
Upper Georgia Strait	5.83%	4.57%	54.85%	20.38%	Upper Georgia Strait
Fraser Early	3.65%	3.96%	24.45%	5.14%	Upper Fraser Middle Fraser Thompson
Columbia Upriver Summer	7.25%	3.89%	26.90%	15.60%	Columbia Upriver Summer
Alaska South SE	1.63%	3.84%	66.63%	55.10%	King Salmon Andrew Creek Blossom Keta Unuk Chickamin
WCVI Wild	1.37%	2.97%	54.14%	18.21%	WCVI
Washington Coastal Wild	1.75%	2.77%	37.61%	9.28%	Grays Harbor Fall Quillayute Fall Hoh Fall Queets Fall
WA Coastal Hatchery	1.80%	2.32%	17.03%	3.71%	NA
Willamette River Hatchery	1.92%	1.24%	12.86%	5.45%	NA
Fall Cowlitz Hatchery	0.23%	1.05%	4.35%	2.27%	NA
Lewis River Wild	1.26%	0.89%	17.67%	8.69%	Lewis River
Lower BC Hatchery	0.20%	0.33%	3.64%	1.82%	NA
PS Hatchery Fingerling	0.13%	0.22%	0.47%	0.29%	NA
Lower Georgia Strait	0.20%	0.20%	2.97%	2.07%	Lower Georgia Strait
Fraser Late	0.19%	0.15%	0.31%	0.11%	Harrison
Snake River Fall	0.29%	0.12%	6.42%	4.04%	Not Represented
Spring Cowlitz Hatchery	0.17%	0.10%	1.96%	1.05%	NA
Skagit Summer/Fall	0.02%	0.10%	4.46%	1.20%	Skagit Summer/Fall
Stillaguamish Summer/Fall	0.00%	0.08%	10.75%	7.00%	Stillaguamish
PS Yearling	0.06%	0.06%	0.61%	0.56%	NA
Nooksack Fall	0.02%	0.04%	0.15%	0.14%	NA
Puget Sound Natural	0.07%	0.04%	0.71%	[illegible]	Green
Snohomish Summer/Fall	0.02%	0.02%	6.80%	1.20%	Snohomish
Spring Creek Hatchery	0.00%	0.00%	0.00%	0.00%	NA
Lower Bonneville Hatchery	0.00%	0.00%	0.00%	0.00%	NA
Nooksack Spring	0.00%	0.00%	0.00%	0.00%	Not Represented

6

SEAK Stock Composition Example

FISHERY:	SE ALASKA ALL GEAR				
	2015	Average (1985-2014)			
Model Stock	% of Fishery Catch	% of Fishery Catch	% of Stock Catch	% of Stock Total Return	Associated Escapement Indicator Stocks
Columbia Upriver Bright	30.97	18.09	25.51	13.22	Columbia Upriver Bright
WCVI Hatchery	10.66	15.50	52.64	18.12	NA
Oregon Coastal North Migrating	7.75	13.46	34.01	15.14	Oregon Coastal
Columbia Upriver Summer	7.25	3.89	36.90	15.60	Columbia Upriver Summer
WCVI Wild	1.37	2.97	54.14	18.21	WCVI (14 stocks)
Spring Creek Hatchery	0.00	0.00	0.00	0.00	NA

7

Proportion of Model Stock in SEAK Troll

8

Stock Distributions

Mean % total mortality in PSC fisheries and escapement for CWT ERA stocks & associated natural escapement indicators during the 2009 PST and % change relative to 1999 PST.

Escapement Indicator Stock	Expl. Rate Ind.	AABM						ISBM				ESC	
		SEAK		NBC		WCVI		US		CAN			
		% Mort	% Chg	% Mort	% Chg	% Mort	% Chg	% Mort	% Chg	% Mort	% Chg	% Esc	% Chg
Col R Summer	SUM	12%	-7%	3%	-3%	9%	-3%	41%	19%	1%	0%	35%	-6%
Col Upriver Bright	URB	13%	-4%	5%	-1%	3%	0%	32%	6%	1%	0%	47%	-2%
Nehalem Siletz Siuslaw	SRH	14%	-6%	10%	0%	2%	1%	30% Max	3%	0%	0%	43%	2%
Spring Creek	SPR	0%	0%	0%	0%	7%	-6%	60%	14%	1%	1%	32%	-8%
WCVI Hatchery	RBT	18%	1%	6%	0%	3%	2%	0%	0%	25% Max	-5%	48%	1%
WCVI Adjusted		19%	0%	6%	0%	3%	2%	0%	0%	9% Min	3%	63%	-5%

Note: "WCVI Adjusted" row shows effect of removing mortalities occurring in terminal fisheries targeting the Robertson Creek hatchery CWT indicator stock; the WCVI 2 rows bound impacts on the natural stocks in the aggregate. 9

PSC Chinook Model – John Carlile (ADFG) & Antonio Velez-Espino (CDFO)

PSC Chinook Model: The Purpose, the Structure and the Use of Forecasts in the Model

John K. Carlile, Alaska Department of Fish and Game
Antonio Velez-Espino, Canadian Department of Fisheries and Oceans
Gayle Brown, Canadian Department of Fisheries and Oceans

1

Chinook Model Structure and Assumptions

2

Chinook Model Basic Facts

- Deterministic approach
- Annual time periods
- Single pool (i.e., no explicit migration)
- Incorporates data from CWT-based cohort analyses
- Incorporates historical data (e.g., catch, terminal run/escapement)
- Scales abundance to ERs from a base period (1979-82)
- Currently 25 fisheries & 30 stocks (48 fisheries & 40 stocks in the future)
- 1999- Model changed from a rebuilding assessment tool to a management tool

3

Main Chinook Model Assumptions

- Ocean distribution of individual stocks is consistent with those experienced during the model base period (static)
- Hatchery indicator stocks are reasonable surrogates for wild stocks in the same geographic area with similar life histories (i.e., age structure, maturation, ocean distribution)
- All stocks of a given age have the same size distribution in a given fishery

4

Alternative Models

- Model using continuous catch equations proposed in 2004 (Gary Morishima, Din-Geng Chen), Funded by US LOA
- Advantages – could better account for interactions between fisheries, temporal stratification of fisheries easier, provide more information on the variability of stock distributions
- Disadvantage – need estimates of effort for each fishery which aren't readily available for some fisheries
- Catch at Age Model proposed in 2004 (Rishi Sharma and Henry Yuen), Not funded by US LOA

5

Why Haven't the Alternate Models Been Explored?

- The Pacific Salmon Treaty dictates that management of a number of Chinook fisheries is tied to abundance indices from the PSC Chinook model so there has been a reluctance to modify or replace the model
- A Data Generation Model (DGM) is currently being developed that will allow sample datasets to be generated
- The DGM datasets will allow for comparison of output statistics from different management models against known parameters (cohort sizes, exploitation rates, etc)

6

Stocks and Fisheries in the Chinook Model

STOCK #	STOCK	FISHERY #	FISHERY
1	**Southeast Alaska**	1	**Alaska Troll**
2	North/Central BC	2	North BC Troll
3	Fraser Early	3	Central BC Troll
4	Fraser Late	4	WCVI Troll
5	West Coast Vancouver Island Hatchery	5	WA/OR Troll
6	West Coast Vancouver Island Natural	6	Georgia Strait Troll
7	Georgia Strait Upper	7	**Alaska Net**
8	Georgia Strait Lower Natural	8	North BC Net
9	Georgia Strait Lower Hatchery	9	Central BC Net
10	Nooksack Fall	10	WCVI Net
11	Puget Sound Hatchery Fingerling	11	Juan De Fuca Net
12	Puget Sound Natural Fingerling	12	Puget North Net
13	Puget Sound Hatchery Yearling	13	Puget South Net
14	Nooksack Spring	14	Washington Coastal Net
15	Skagit Wild	15	Columbia River Net
16	Stillaguamish Wild	16	Johnstone Strait Net
17	Snohomish Wild	17	Fraser Net
18	Washington Coastal Hatchery	18	**Alaska Sport**
19	Columbia UpRiver Brights	19	North/Central BC Sport
20	Spring Creek Hatchery	20	WCVI Sport
21	Lower Bonneville Hatchery	21	WA Ocean Sport
22	Fall Cowlitz Hatchery	22	Puget North Sport
23	Lewis River Wild	23	Puget South Sport
24	Willamette River	24	Georgia Strait Sport
25	Spring Cowlitz Hatchery	25	Terminal (Col Riv) Sport
26	Columbia River Summer		
27	Oregon Coastal		
28	Washington Coastal Wild		
29	Lyons Ferry (Snake River Fall)		
30	Mid-Columbia River Brights		

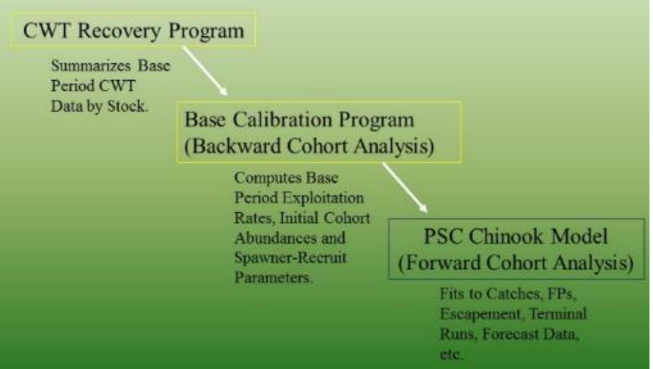
7

Model Input Data

- Base Period CWT Data
- Fishery Catch Data
- Chinook Non-Retention Data
- Past Escapement and/or Terminal Run Data
- Terminal Run/Escapement Forecasts
- Fishery Policy (FP) - Exploitation Rate Scalars
- Maturation Rate and Adult Equivalent Data
- Hatchery Enhancement Data
- Spawner-Recruit Parameters
- Proportion Non-Vulnerable (PNV) Changes
- Inter-Dam Loss Factors

8

Complete Calibration Procedure



9

Data files used by the Chinook Model

- .STK – Base period data for individual stocks (cohort size, maturation rates, AEQs, exploitation rates)
- .BSE – Stocks, ages, fisheries, start year, end year, ocean net age, Fishery names, Natural Mortality Rates, Incidental Mortality Rates, Ricker Stock-Recruit Parameters
- .ENH – enhancement changes
- .IDL – post-fishery, pre-spawning mortality
- .MAT – annual maturation rate estimates from ERA
- .PNV – proportion non-vulnerable for individual fisheries (size limits)
- .CNR – chinook non-retention
- .CEI – catch levels for fisheries
- .FP – fishery policy factors (use ERA harvest rate data)
- .FCS – escapement/terminal run historical data and __forecasts__

The main purpose of the FCS file is EV estimation and EV forecasting

10

Chinook Model Calibration Process

11

Model Calibration Stages

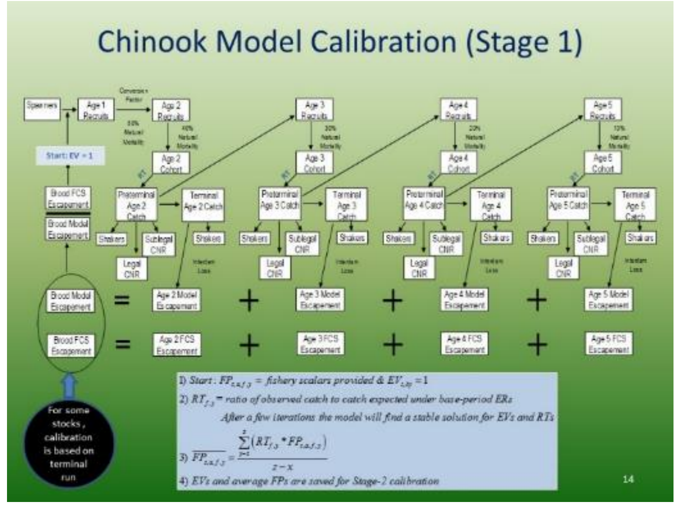
- Stage-1 Calibration
  - Main purpose: estimate preliminary Environmental Variable (EV) factors and generate RT scalars
- Stage-2 Calibration
  - Main purpose: fine tune the EV factors for recent broods
- Projection Run
  - Main purpose: produce the preseason and postseason estimates of abundance indices

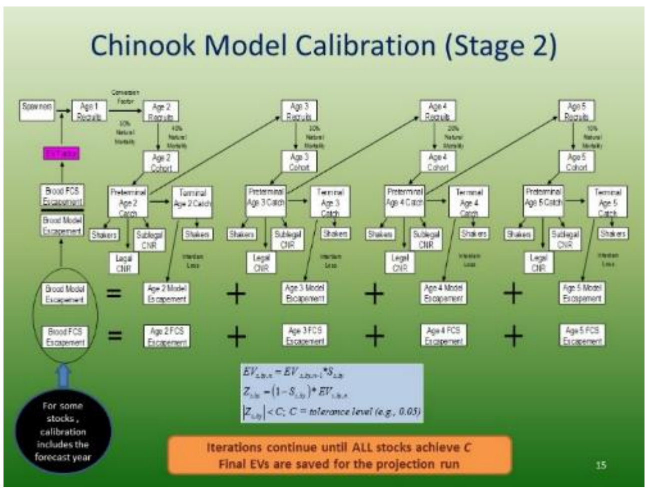
12

EVs

- EV factors scale the number of recruits produced by the stock-specific spawner-recruit parameters to match supplied escapement/terminal run values
- EV factors are stock and brood-year specific
- Thus, age composition must be generated
- Stage-2 calibration generates the EVs that are going to be used for the projection run
- EVs are interrelated (i.e., each iteration will change the EVs for all stocks and brood years)
- If the spawner-recruit parameters are appropriate for a stock the EVs can be thought of as survival scalars. If they are not appropriate then the EVs can be thought of as survival scalars combined with "slop" factors

13

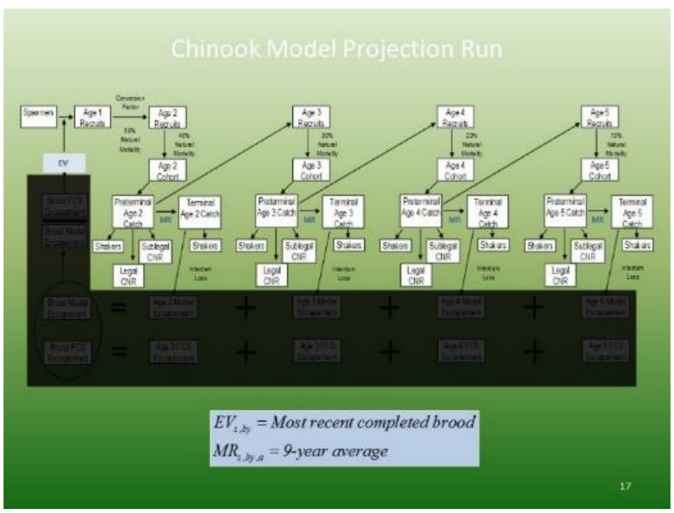


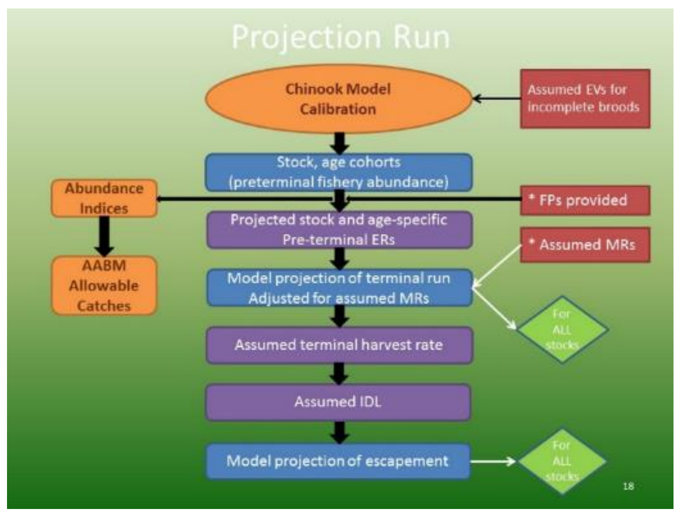


Projection Run (Stage 3)

- The purpose of Stage 1 and Stage 2 is to fit to the brood-year terminal runs/escapements by generating stock and brood year specific EVs
- The purpose of the projection run is to estimate the cohort abundances for the years following the last year with observed data using the EVs from Stage 2

16





The Role of Forecasts in the Chinook Model

19

Forecast (FCS) File Data

- 28 sets of time series in the FCS file
  - 30 stocks represented (RBT+RBH; PSF+PSY)
  - 9 for escapement & 19 for terminal run
- All time series include observed (actual; starting in 1979) data except for the last (forecast) year
- 22/28 usually include data for forecast year
- 16/28 include age-specific data

20

Forecast (FCS) File Data continued…

- Agency forecasts prepared using wide range of methods
- Forecasts needed in March when data from the previous year may not be available (but forecasts produced anyway)
- Forecasts generally used in the Model calibration without scrutiny
- Model stocks range from individual stocks (Nooksack Springs) to large aggregates of many stocks (Fraser Early!)
- Aggregates are mixtures - natural spawning stocks, hatchery stocks

21

Agency and Model Forecasts

- The Chinook model produces a stock forecast (terminal run/escapement) regardless of whether an agency forecast is provided
- No agency forecast – recent EV averages determine the model forecast
- The model calibrates (fits) to the total brood year terminal run/escapement exactly but the calendar year values differ due to uncertainty in estimating the age composition (maturation rate assumptions etc)

22

Agency Forecasts have an effect on the most recent EVs: example from 2015 FCS file

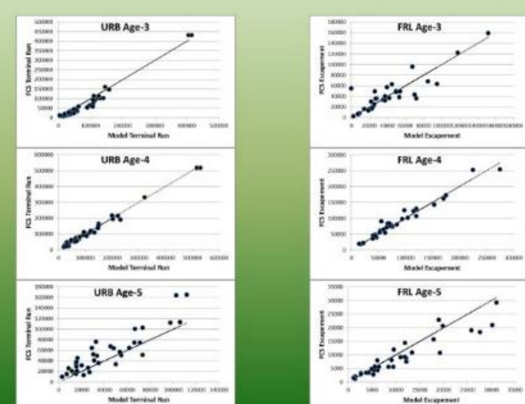
SPR		Spring Creek Tules Bonneville Pool Hatchery					
	2	read in maturity schedule-do not adjust					
	2	Terminal Run Adults					
	3	First age					
	1	Minimum number of ages				Brood Year	
79	1	66500	28000	700		1976	
80	1	57300	40300	200		1977	
81	1	59700	25300	1300		1978	
82	1	94400	26200	200		1979	
83	1	14437	14148	314		1980	
84	1	39234	8115	187		1981	
85	1	24336	8066	832		1982	
86	1	9782	6566	229		1983	
87	1	5500	3500	100		1984	
.	.	.	.	.	.	.	
.	.	.	.	.	.	.	
.	.	.	.	.	.	.	
110	1	125197	5542	29		2007	
111	1	55415	15128	34		2008	
112	1	43089	13606	71		2009	
113	1	69565	16862	142		2010	Influenced by AF
114	1	104526	22192	273		2011	Influenced by AF
115	1	126200	34200	100		2012	Influenced by AF
116	0	1	0				

Last complete brood is 2010 (thanks to forecast data)

Last complete brood would be 2009 without forecast data

23

Example of Model Fit (1979-2014)



24

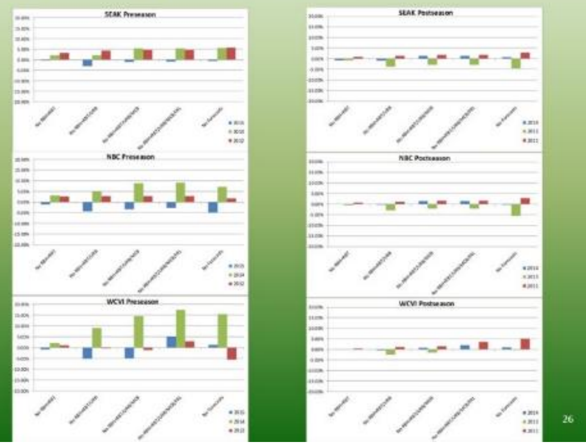
Example: WCVI Forecast Data

WCVI Terminal Runs

Modified Calibration	Year	With Agency Forecast Preseason Agency (Combined H+W) Age 3	Age 4	Age 5	Total	Model Hatchery	Wild	Total	Without Agency Forecast Preseason Model Hatchery	Wild	Total	Post-Season Agency (Combined H+W) Age 3	Age 4	Age 5	Total
CLB0107	2001	21500	3333	5800	30633	30402	6152	36554	17264	3231	20495	61484	16803	8500	86787
CLB0206	2002	38960	64257	6665	109882	96506	16989	113495	133010	23694	156704	53473	111908	2801	168182
CLB0308	2003	19794	53699	32308	105801	88016	9691	97707	113081	11386	124467	48927	134562	31856	215345
CLB0401	2004	53476	54610	36094	144180	135851	8722	144573	157199	9436	166635	115470	71222	60808	247500
CLB0506	2005	45261	135425	38154	218840	196068	13926	209994	303180	23533	326713	27836	107457	19301	154594
CLB0604	2006	54139	33912	50827	138878	121247	8916	130163	106397	8013	114410	86929	87336	22832	197097
CLB0705	2007	6798	89230	21293	117321	120168	13786	133954	219833	26685	246518	5291	100358	12433	118082
CLB0807	2008	15441	6168	38646	60255	47003	5815	52818	38645	4780	43425	46092	24380	28272	98744
CLB0907	2009	9148	43520	5714	58382	56050	6985	63035	113571	14765	128336	24917	55809	7703	88429
CLB1007	2010	19840	23470	18276	61586	53616	7226	60842	65386	8994	74380	62067	23936	6531	92534
CLB1106	2011	6538	58821	9349	74708	73475	9181	82656	146908	17189	164097	16337	138680	6897	161914
CLB1209	2012	2870	15108	36787	54765	46252	5469	51721	63985	7182	71167	6944	54183	23305	84432
CLB1308	2013					27153	2802	29955	27152	2802	29954	79413	72448	21771	173632
CLB1402	2014	12693	176739	27295	216727	166201	20530	186731	211536	27830	239366	17546	92350	10577	120473
CLB1503	2015	43624	12921	48438	105003	79820	9806	89626	65030	8605	73635	112455	57301	20008	189764
CLB1601	2016	26186	185708	12225	224119	192626	21981	214607	283591	34228	317819				

25

Sensitivity of the AIs to Agency Forecast Data



26

What Determines the Effect of a Stock's Forecast on Fishery AIs?

- The proportion of the fishery that stock represents (major contributors have more of an effect on the AI)
- The magnitude of the difference in prefishery cohort sizes due to model's EVs produced with and without agency forecast data
- The accuracy (and age composition) of agency forecasts for ALL stocks
- Interactions with other input data (e.g., FPs, MRs, etc.)

27

Model Abundance Indices (and other outputs)

28

Model Outputs

(Not exhaustive)

- Catches by fishery, stock and age
- Incidental mortalities by fishery, stock and age
- Fishery specific stock composition estimates
- Exploitation rates by fishery, stock and age
- Terminal runs/escapements by stock and age (original intent of the model)
- Abundance Indices for Southeast Alaska, Northern British Columbia and West Coast of Vancouver Island (current focus of the model)

29

Calculation of the Abundance Index

$$AI_{fy} = \frac{\sum_s \sum_a \left[C_{sa(b=y-a)} * SR_a * ER_{fsa} * (1 - PNV_{fa}) \right]}{\sum_{1979}^{1982} \left\{ \sum_s \sum_a \left[C_{sa(b=y-a)} * SR_a * ER_{fsa} * (1 - PNV_{fa}) \right] \right\} / 4}$$

Where

AI = Abundance Index
C = Cohort Size
SR = Survival Rate
ER = Base Period Exploitation Rate
PNV = Proportion Non-Vulnerable
f = fishery
s = stock
a = age
b = brood year
y = calendar year

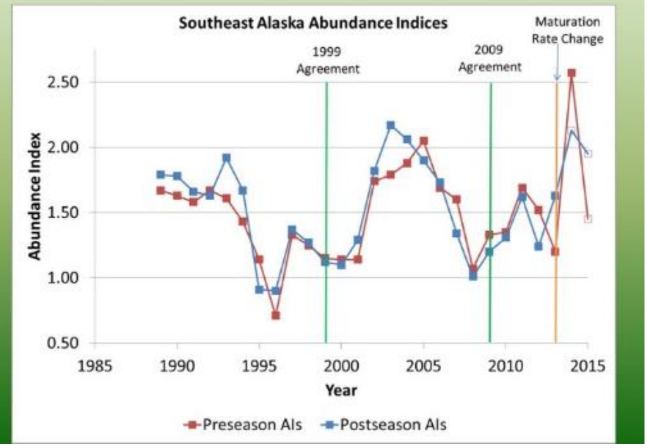
30

AIs from the AABM Fisheries

	SEAK		NBC		WCVI	
Year	Preseason	Postseason	Preseason	Postseason	Preseason	Postseason
1999	1.15	1.12	1.12	0.97	0.60	0.50
2000	1.14	1.10	1.00	0.95	0.54	0.47
2001	1.14	1.29	1.02	1.22	0.66	0.68
2002	1.74	1.82	1.45	1.63	0.95	0.92
2003	1.79	2.17	1.48	1.90	0.85	1.10
2004	1.88	2.06	1.67	1.83	0.90	0.98
2005	2.05	1.90	1.69	1.65	0.88	0.84
2006	1.69	1.73	1.53	1.50	0.75	0.68
2007	1.60	1.34	1.35	1.10	0.67	0.57
2008	1.07	1.01	0.96	0.93	0.76	0.64
2009	1.33	1.20	1.10	1.07	0.72	0.61
2010	1.35	1.31	1.17	1.23	0.96	0.95
2011	1.69	1.62	1.38	1.41	1.15	0.90
2012	1.52	1.24	1.32	1.15	0.89	0.76
2013	1.20	1.63	1.10	1.51	0.77	1.04
2014	2.57	2.20	1.99	1.80	1.20	1.12
2015	1.45	1.95	1.23	1.69	0.85	1.05
2016	2.06		1.70		0.89	

31

How well does it work?



Maturation Rates and Their Effect on the AIs

33

Maturation Rate Issues (Part 1)

- Pre-2013 Model assumed long-term average maturation rates for recent incomplete broods
- Discovered that a number of stocks showed trend toward maturing at younger ages [TCCHINOOK(16)-1]
- AIs are calculated on vulnerable cohorts, younger fish are less vulnerable meaning lower abundance but the model assuming older fish
- Fix implemented in 2013 model calibration, replace long-term average maturation rates with recent 5-year average maturation rates

34

Maturation Rate Issues (Part 2)

- 2013 change in maturation rate average coincided with an end the chronic over-prediction of the preseason AIs (still too early to determine if the bias is gone for good) BUT…
- 2013-2015, large discrepancies between pre and post-season AIs and possible cyclical behavior
- 2016, revisited the maturation rate average and determined that a nine-year recent average would perform better than a five-year average

35

Recent Controversy

- What is the cause of the large discrepancies between the pre and post-season AIs in recent years?
- How much is due to inaccurate terminal run/escapement forecasts provided by the agencies?
- How should agency forecasts be reviewed (e.g. this workshop)?
- No CTC consensus was reached on 2015 or 2016 Model calibrations (settled by the Commission)

36

Discussion of the PSC Chinook Model – Central Questions for the Panel's Focus

Table 4. Central issues or improvements identified by workshop participants for consideration by the panel with respect to the PSC Chinook Model.

Incorporate uncertainty directly
Accommodation of changes in hatchery smolt releases over time
Changes in fish distribution among fisheries over time (not constant)
Forecasts are not available for all stocks
Data vs. assumptions
Application of base period to fishery policy inputs
Hatchery vs. wild treatment/assumption
Pattern of staff succession, transfer of institutional knowledge, documentation
Limited CTC capacity relative to Treaty reporting demands (also related to the paucity of work on model innovation)
Computer programming support within PSC for analytical modernization
Model evaluation criteria
Availability of pre and in-season information for recognition of outlier years
Incorporation of environmental information for improving forecasts
Opportunities for new model construction/application (e.g. Morishima, Sharma)
Do nothing (recent anomalies are just atypical years)
Strain of adding model complexity & stratification
Clarify objectives of competing model uses (abundance index v. exploitation rate by stock)
Model simplification (e.g. simple weighted average of forecast abundance
Simplification of fishery management regimes

West Coast Vancouver Island Forecast – Diana Dobson (CDFO)

Fisheries and Oceans Canada   Pêches et Océans Canada

WCVI Chinook Forecast: Review of Methods and Performance

Diana Dobson, DFO

Lead Biologist,

WCVI Salmon Stock Assessment

Canada

Fisheries and Oceans Canada   Pêches et Océans Canada

WCVI Chinook Forecast

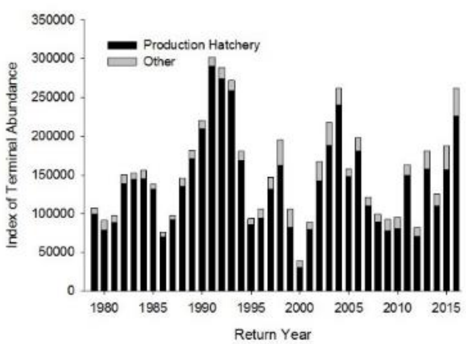
- For the CTC model, the age 3, 4 and 5 year old **terminal return of chinook** is forecast for WCVI index stocks (3 major hatcheries + 18 index stocks, many of which are enhanced).
- This terminal index likely accounts for >95% of WCVI chinook production. Averages about 150K return (range 40 to 300K)
- The terminal index from run reconstruction methods; input data is variable across stocks/areas.
- Average terminal age composition of WCVI chinook: Age 3 – 30%, Age 4 – 50%, Age 5 – 20%

2

Canada

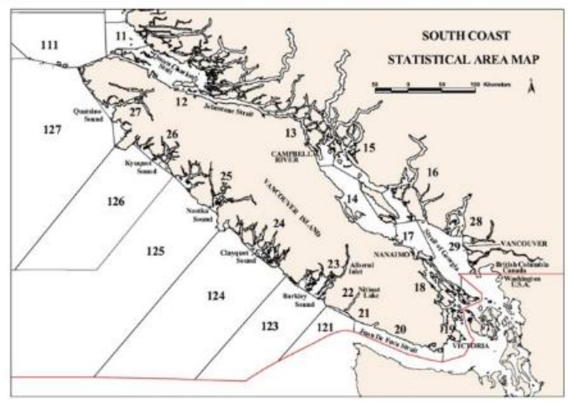
Fisheries and Oceans Canada   Pêches et Océans Canada

Terminal Index, 1979 - 2015



Canada

Fisheries and Oceans Canada   Pêches et Océans Canada



4

Canada

Fisheries and Oceans Canada   Pêches et Océans Canada

Forecast Performance

1. RCH CWT Production
2. Somass/RCH Terminal Run (Domestic Management Forecast)

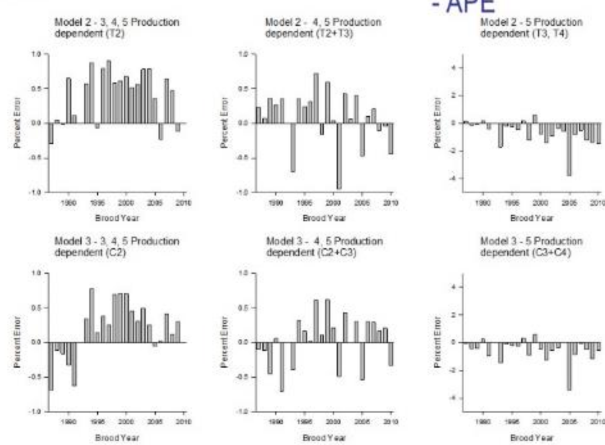
- ➢ Forecast evaluated using the data to t-1
- ➢ Percentage error plotted for forecast year + tabulated other summary statistics for precision, bias.

5

Canada

Fisheries and Oceans Canada   Pêches et Océans Canada

RCH CWT Production Forecast - APE



Fisheries and Oceans Canada Pêches et Océans Canada

Dependent		RCH CWT Ass. Production		
		3,4,5	4,5	5
Model 2	Independent	Terminal 2	Terminal 3	Terminal 5
	MAD	36,353	18,710	6,755
	MAPE	49%	33%	78%
	Mean Percent Error	42%	8%	-72%
	Median Est	75%	67%	17%
Model 3	Independent	Prod 2	Prod 3,4	Prod 5
	MAD	35,112	16,380	6,707
	MAPE	36%	30%	66%
	Mean Percent Error	20%	3%	-57%
	Median Est	71%	63%	17%
AVERAGE Model	MAD	**32,683**	**13,007**	**6,562**
	MAPE	**39%**	**27%**	**74%**
	Mean Percent Error	**31%**	**6%**	**-64%**
	Median Est	**75%**	**63%**	**17%**
AVG OBS	**1987-2010**	**88,386**	**61,555**	**15,284**

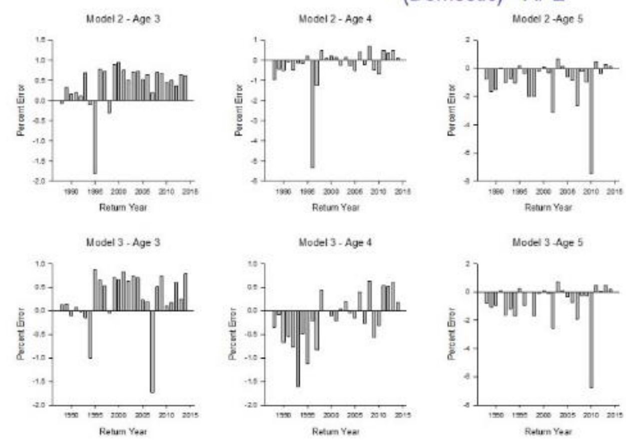
RCH CWT Production Forecast

7

Canada

Fisheries and Oceans Canada Pêches et Océans Canada

Somass/RCH Terminal Forecast (Domestic) - APE



Fisheries and Oceans Canada Pêches et Océans Canada

Somass/RCH Terminal Forecast (Domestic Management)

Dependent		Terminal Somass/RCH Run			
		Age 3	Age 4	Age 5	Total
Model 2	Independent	Terminal 2	Terminal 2,3	Terminal 4,5	
	MAD	13,358	17,118	9,795	**25,609**
	MAPE	56%	58%	110%	**30%**
	Mean Percent Error	39%	-38%	-95%	**4%**
	Median Est	86%	46%	29%	**54%**
Model 3	Independent	Prod 2	Prod 3,4	Prod 5	
	MAD	10,574	23,949	10,818	**38,068**
	MAPE	50%	45%	104%	**37%**
	Mean Percent Error	28%	-15%	-87%	**0%**
	Median Est	79%	39%	32%	**50%**
AVERAGE Model	MAD	11,330	22,880	10,001	**33,978**
	MAPE	48%	49%	116%	**33%**
	Mean Percent Error	34%	-21%	-101%	**3%**
	Median Est	86%	39%	29%	**54%**
AVG OBS	**1988-2015**	**29,242**	**44,731**	**14,606**	**88,580**

9

Canada

Fisheries and Oceans Canada Pêches et Océans Canada

Forecast performance v. Model Input Error

- What is the purpose of the terminal WCVI forecast as input to the CTC model?
- Need to distinguish between the CTC model input requirement and terminal forecasts that are generated for domestic fishery management:

$$AI_{f,y} = \frac{\sum_s \sum_a COH_{a,y} * ER_{a,f} * (1-PNV_{a,f})}{\sum_s \sum_a COH_{a,b} * ER_{a,f} * (1-PNV_{a,f})}$$

Where:

COHs,a,y= Cohort size for stock ***s***, age ***a***, in year ***y***

COHs,a,b= Average Cohort size for stock ***s***, age ***a***, during 1979-1982

ERs,a,f = 1979-1982 average base period exploitation Rate for stock ***s***, age ***a***, in fishery ***f***

PNVa,f = Proportion of age ***a*** fish that is not vulnerable in fishery ***f*** under regulations in effect during a specified period, usually 1979-1982

10

Canada

Fisheries and Oceans Canada Pêches et Océans Canada

Forecasting Method

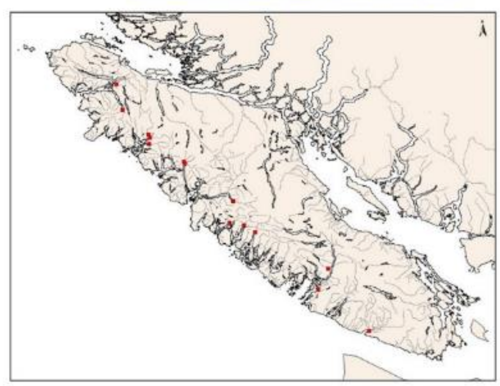
1. The basic forecast method uses a 'sibling' model to predict the production of older age classes from the observed production of younger age classes from the same brood using simple linear regressions. (The regressions are developed for CWT associated production from the Robertson Creek Hatchery (RCH) Indicator Stock).
2. The forecast CWT associated production for RCH is expanded for the entire Somass/RCH system.
3. The terminal return of Somass/RCH is then predicted after applying assumptions for 'pre-terminal' fishing mortality, stock composition in pre-terminal fisheries, and maturation rates.
4. The forecast terminal return of Somass/RCH is then expanded for the WCVI index systems.

11

Canada

Fisheries and Oceans Canada Pêches et Océans Canada

WCVI Management Unit



Canada

Fisheries and Oceans Canada | Pêches et Océans Canada

For the Robertson Creek Hatchery (RCH) CWT Indicator Stock, simple linear regressions are developed from <u>production data</u> generated from the CTC cohort analysis*:

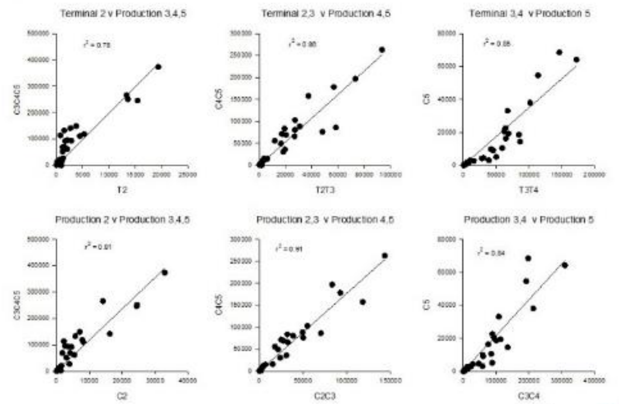
- Model 2 (Prod2) uses total terminal return at a younger age class (independent variable) to predict total production (the surviving cohort in the ocean) of a subsequent age or ages from the same brood year. The dependent variable is the total (total ocean fishing mortality plus terminal run) production at a subsequent age or ages.
- Model 3 (Prod3) model uses estimated total production (total fishing mortality plus escapement) of an age class(es) to predict total production of subsequent ages (i.e., the surviving cohort, dependent from the same brood year).

The cohort analysis makes assumptions regarding natural mortality, maturation rate .

13

Canada

Fisheries and Oceans Canada | Pêches et Océans Canada



14

Canada

Fisheries and Oceans Canada | Pêches et Océans Canada

The forecast for the CWT associated production for Robertson Creek Hatchery is then expanded for the entire Somass/RCH system based on ratios of earlier returns from the brood year:

Return Year	Terminal RCH CWT Return				Total Somass/RCH Return			
	Age 2	Age 3	Age 4	Age 5	Age 2	Age 3	Age 4	Age 5
2014	6,091	2,705	18,189	1,076	7,023	3,799	21,161	1,283
2015	1,551	29,540	1,468	553	1,887	64,458	5,551	1,167

$$\text{Age 3, 4, 5 Expansion} = \frac{(Total\ Age\ 2\ SomassRCH\ return)_{t-1}}{(RCH\ Age\ 2\ CWT\ return)_{t-1}}$$

$$\text{Age 4,5 Expansion} = \frac{(Total\ Age\ 3\ SomassRCH\ return)_{t-1}}{(RCH\ Age\ 3\ CWT\ return)_{t-1}}$$

$$\text{Age 5 Expansion} = \frac{((Total\ Age\ 4\ SomassRCH\ return)_{t-1} + (Total\ Age\ 3\ SomassRCH\ return)_{t-2})}{((RCH\ Age\ 4\ CWT\ return)_{t-1} + (RCH\ Age\ 3\ CWT\ return)_{t-2})}$$

15

Canada

Fisheries and Oceans Canada | Pêches et Océans Canada

After the Somass/RCH production for each brood year is forecast, assumptions are then applied to predict the terminal run size of Somass/Robertson Creek Hatchery chinook.

These assumptions relate to:

1. Pre-terminal fishery mortality;
2. Pre-terminal fishery stock composition*; and
3. Maturation rate*.

** both generated by the cohort analysis*

16

Canada

Fisheries and Oceans Canada | Pêches et Océans Canada

Pre-terminal fishery mortality

- Pre-terminal fishery mortality needs to be predicted for each age class/brood year.
- There is some uncertainty as to what is the correct assumption for CTC model calibration purposes:
  - *Is it the pre-terminal fishery mortality that was exerted during the base period? Or,*
  - *Is it the expected pre-terminal fishery mortality given current fishing regimes?*
- In recent years, both options have been presented although the latter has been applied for the input of WCVI terminal run for the CTC calibration.

17

Canada

Fisheries and Oceans Canada | Pêches et Océans Canada

2016 FORECAST

Assumption	BASE-Period Fishery		fcs
Forecast	Pre-fishery Somass/RCH	Terminal Somass/RCH	WCVI Terminal
2013 brood	48,671	6,066	*22,716*
2012 brood	249,518	86,611	*174,555*
2011 brood	2,993	1,515	*14,606*
Total	301,182	94,192	*211,878*

Forecast terminal WCVI CN return, **base-period units***

2016 FORECAST

Assumption	3 YR AVG Mortality		fcs
Forecast	Pre-fishery Somass/RCH	Terminal Somass/RCH	WCVI Terminal
2013 brood	48,671	6,992	*26,186*
2012 brood	249,518	92,144	*185,708*
2011 brood	2,993	1,268	*12,225*
Total	301,182	100,405	*224,119*

Forecast terminal WCVI return, **like recent year fisheries***

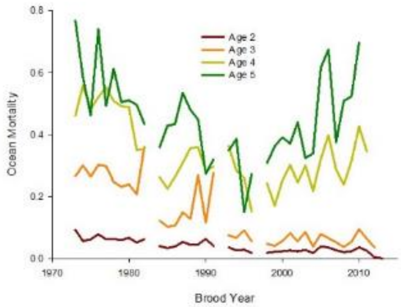
** Note: both forecasts are different from the 'domestic' management forecasts, which are only completed after CTC calibration and AABM fishery Abundance Indices/TACs are known.*

18

Canada

Fisheries and Oceans Canada / Pêches et Océans Canada

Age-specific RCH pre-terminal exploitation rate



Canada

Fisheries and Oceans Canada / Pêches et Océans Canada

WCVI Index Expansion

The Somass/RCH terminal forecast was expanded for the WCVI index stocks using a similar method as applied for the expansion from RCH CWT production to Somass/RCH Total production. That is, based on ratios of earlier returning years from the brood:

$$\text{Age 3 Expansion} = \sum_{n=1}^{n} \frac{\textit{WCVI CN Index}}{\textit{Total Somass RCH Return}}/n$$

$$\text{Age 4 Expansion} = \frac{(\textit{Age 3 WCVI Index})_{t-1}}{(\textit{Total Somass RCH Age 3 return})_{t-1}}$$

$$\text{Age 5 Expansion} = \frac{((\textit{Age 4 WCVI Index})_{t-1} + (\textit{Age 3 WCVI index})_{t-2})}{((\textit{Total Somass RCH Age 4 CWT return})_{t-1} + (\textit{Total Somass RCH Age 3 CWT return})_{t-2}))}$$

Canada

Fisheries and Oceans Canada / Pêches et Océans Canada

WCVI Index Expansion

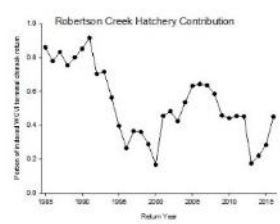
- In more recent years, the Somass/RCH terminal forecast has also been expanded for the WCVI index stocks by adding terminal forecasts that are generated separately for Conuma and Nitinat Hatchery returns and the 18 other index stocks (combined).
- The forecasts generated for the other stocks use information from the RCH CWT cohort analysis (i.e. estimated brood year survival rate) and similar pre-terminal fishery assumptions, but are modified with stock-specific production and age data .

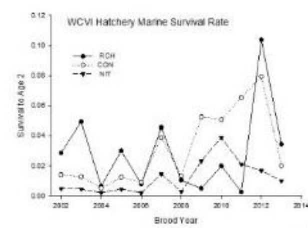
Canada

Fisheries and Oceans Canada / Pêches et Océans Canada

Indicator Stock Assumption

- The objective for developing stock specific WCVI forecasts was to better inform Canadian domestic fishery management. However, variation in stock production/productivity also affects on the WCVI terminal run size forecasts:





Canada

Fisheries and Oceans Canada / Pêches et Océans Canada

Input Data / Years

- All years of RCH CWT cohort data are incorporated in the sibling regressions that form the basis of the WCVI forecast (BY 1983+) – estimates of production, maturation rates, etc.
- Similarly, the WCVI terminal run index has been reconstructed from return year 1979+. All data are incorporated into the forecast and analysis.
- The more challenging issues relate to the varying quality of available assessment data across WCVI systems and the robustness of the indicator stock assumption (e.g. requires similar maturation rate, exploitation patterns, etc.).

Canada

Fisheries and Oceans Canada / Pêches et Océans Canada

Key assumptions, limitations

- Indicator stock – survival, maturation, distribution and exploitation patterns similar among WCVI stocks.
- Generally paucity of data for WCVI stocks other than the RCH CWT indicator and, in some WCVI areas, sample data from fisheries.
- Age 2 recoveries typically low.
- Known bias in CWT recovery data.

Canada

Fisheries and Oceans Canada | Pêches et Océans Canada

Other methods

- Stock-specific forecasts generated (already described)
- Exploration of alternate indices of WCVI terminal run size – e.g. distant fishery indices using additional mark data (DNA, otolith marks)

25

Canada

Fisheries and Oceans Canada | Pêches et Océans Canada

Suggestions for Improvement

- Resolve what input is required for calibration purposes – build a common understanding of the FCS requirement.
- Related - succession, documentation requirement for the process in general – misunderstanding or mis-communication of objectives and/or structural modifications could be a source of error.
- Age 3 forecast – a clear structural issue, also an input problem (age 2 data), and also likely related to changing maturation rate – could explore 'leading indicators' as an alternative method for Age 3 forecasting.

26

Canada

Fisheries and Oceans Canada | Pêches et Océans Canada

Suggestions for Improvement

- Input data – not all information is being used – e.g. available mark data and technology such as DNA, otolith marks, etc.
- Incorporation of uncertainty – not just in reporting of input forecasts, but into the entire management and assessment framework for chinook.
- Simplification of the assessment and management framework – currently data intensive, assumption laden, and resource limited.
- Separation of hatchery from wild abundance in the AIs.

27

Canada

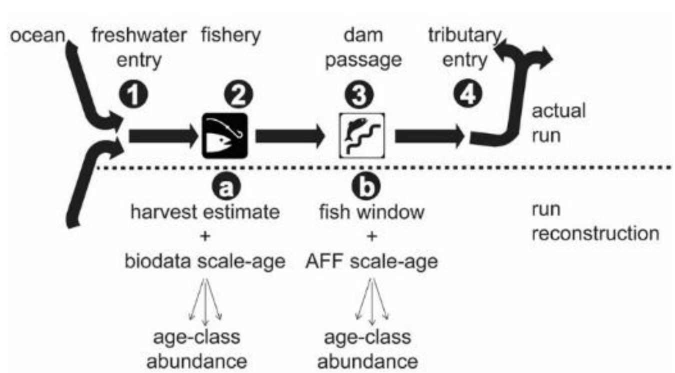
U.S. v. Oregon
Management Agreement

- *United States v. Oregon* is the on-going federal court proceeding that enforces and implements the Columbia River treaty tribes' reserved fishing rights.
- The Parties to *U.S. v. Oregon* are the states of Washington, Oregon, and Idaho; the United States (represented by NOAA Fisheries, the BIA, and USFWS); each of the Columbia River Treaty Tribes (the Confederated Tribes of the Warm Springs Reservation of Oregon, the Confederated Tribes of the Umatilla Indian Reservation, the Nez Perce Tribe, and the Confederated Tribes and Bands of the Yakama Nation); and the Shoshone-Bannock Tribes.
- Fisheries in the Columbia River are managed subject to provisions of *U.S. v. Oregon* under the continuing jurisdiction of the federal court. The 2008-2017 *U.S. v. Oregon* Management Agreement provides the current framework for managing fisheries and hatchery programs in much of the Columbia River Basin.

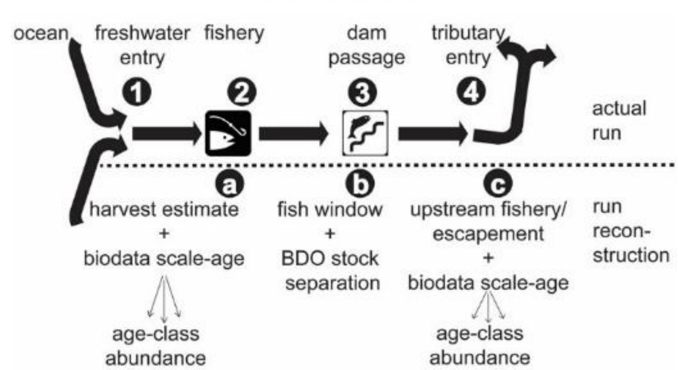
U.S. v. Oregon
Technical Advisory Committee

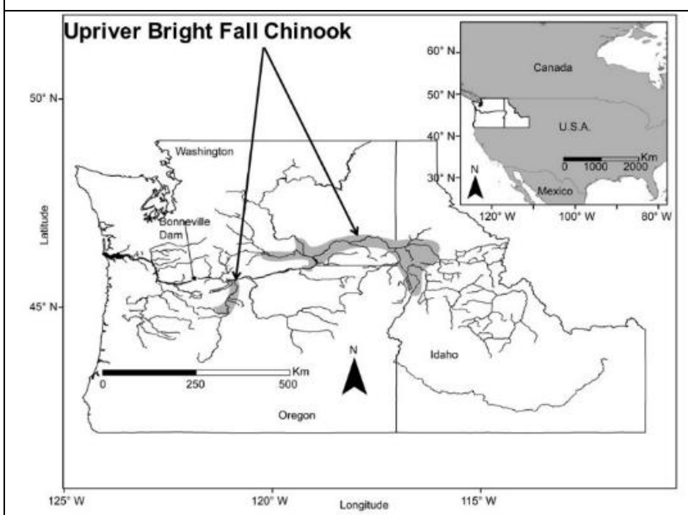
- The *U.S. v. Oregon* Management Agreement establishes a Technical Advisory Committee (TAC) consisting of technical representatives from the entities comprising the Parties to the Agreement.
- TAC was established for the express purpose of developing, analyzing, and reviewing data pertinent to the *US v. Oregon* Management Agreement and to make reports and technical recommendations regarding harvest management.
- TAC reconstructs Columbia River salmon and steelhead returns post-season and develops pre-season forecasts.
    - Either TAC alone, or in coordination with state agency staff (e.g. fall Chinook stocks)
- TAC reviews salmon and steelhead abundance as the runs progress and provides in-season run size updates for stocks originating upstream of Bonneville Dam. The in-season updates allow the States and Tribes to manage fisheries as necessary in order to remain within ESA limits and management guidelines.

Upper Columbia Summer Chinook



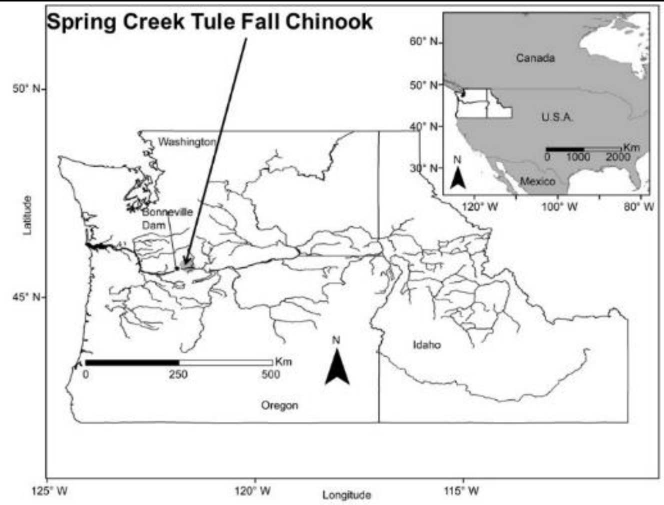
Upriver Bright and Spring Creek Tule Fall Chinook





Upriver Bright (URB)

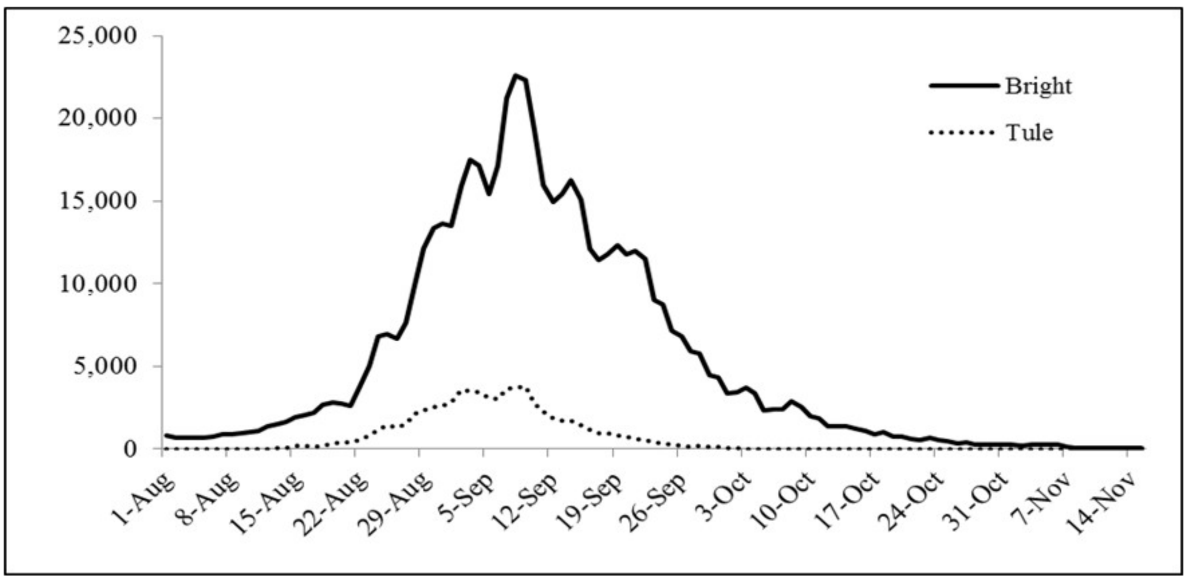
- Bright Stock
- Far-north migrating
- Primarily sub-yearling, ocean-type
- Wild fish represent majority of return
- Majority destined for Hanford Reach, Priest Rapids Hatchery, areas upstream of Priest Rapids Dam, and Snake River.
    - Minor return to Deschutes and Yakima Rivers
- 10-yr average return is nearly 400,000 fish (range 114,500 – 795,900)
    - Past three years have been record returns



Bonneville Pool Hatchery(BPH)
aka Spring Creek Hatchery

- Tule Stock
- Range extends north to British Columbia
- All sub-yearling, ocean-type
- Primarily hatchery fish; some natural production
- Destined for terminal areas within Bonneville Pool
- 10-yr average return is approximately 82,000 fish (range 14,600 – 166,400)
  - Past three years have been strong returns
- Timing at BON more compressed than URBs

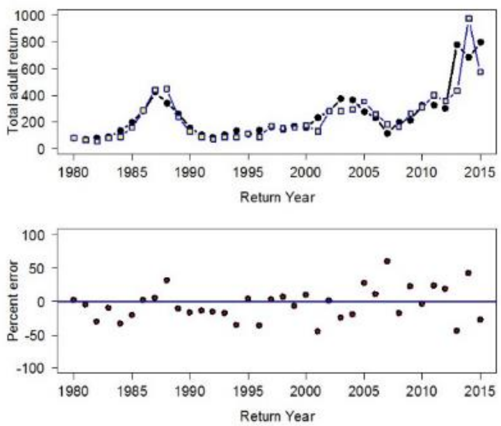
Bonneville Dam Passage Tule v Brights



Average Run Timing of Fall Chinook over Bonneville Dam, applied to the 2016 Forecast.

Columbia River Upriver Bright Forecast - Steve Haeseker (USFWS)

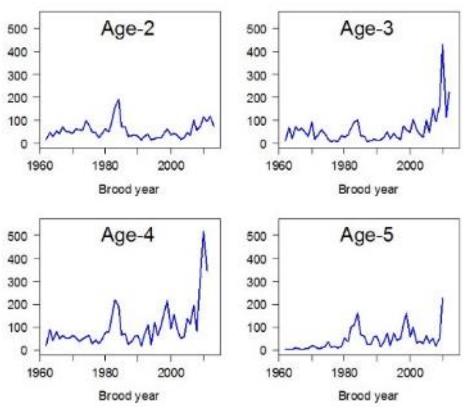
Upriver Bright fall Chinook salmon



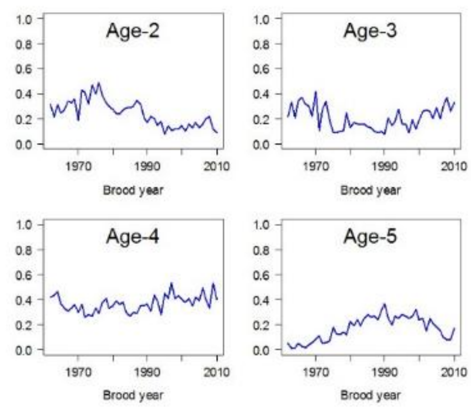
Upriver Bright fall Chinook salmon forecast summary statistics

Forecast source and timeframe	Mean Return	MRE	MPE	MAE	MAPE
TAC 1980-2015	246,300	-11,000	-4.8%	56,100	19.6%
TAC 1999-2013	288,400	-21,300	0.3%	70,000	22.6%
CTC 1999-2013	288,400	-28,500	-1.2%	74,000	23.7%

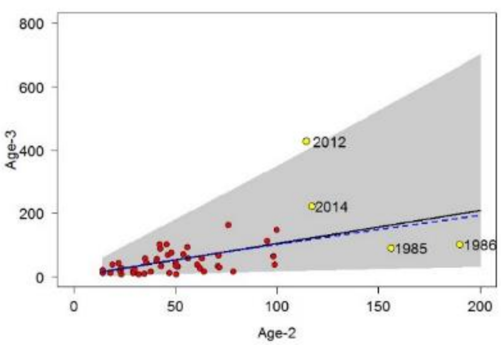
Upriver bright returns by age-class (BY 1962-2013)



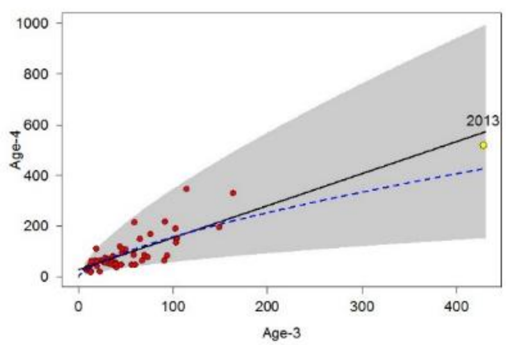
Upriver bright age-class proportions (BY 1962-2010)



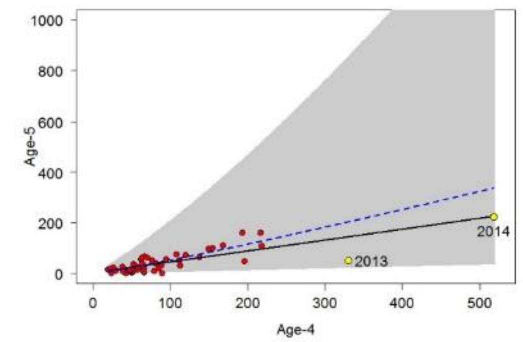
Upriver bright sibling relationships (BY 1962-2012)



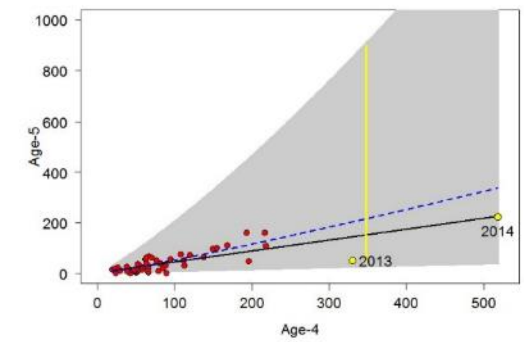
Upriver bright sibling relationships (BY 1962-2012)



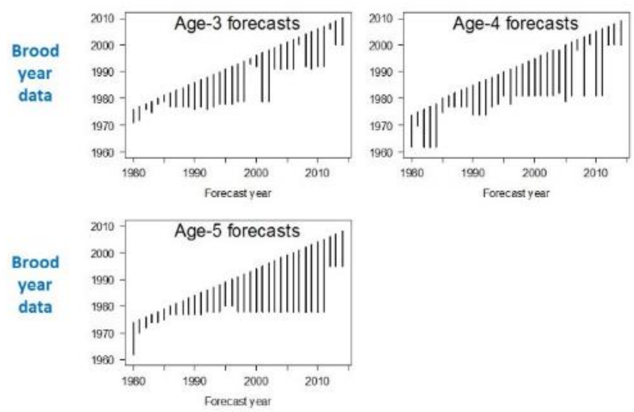
Upriver bright sibling relationships (BY 1962-2012)



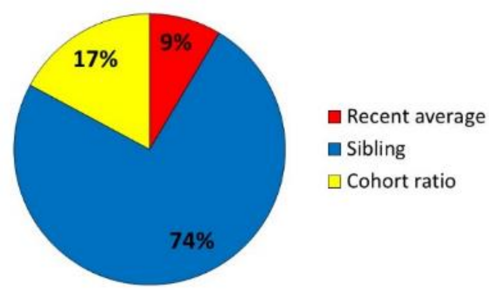
Upriver bright sibling relationships (BY 1962-2012)



Data used for Upriver Bright forecasts



Models selected for forecasting age-3, -4, and -5 returns; (1980-2015)

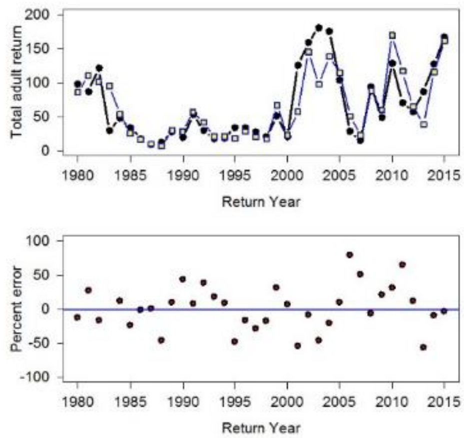


Upriver Bright conclusions

- Natural variability in age-composition makes forecasting challenging
- High quality data on abundance and age-composition
- Forecasts have been unbiased, and demonstrated good precision

Columbia River Spring Creek Hatchery Forecast - Steve Haeseker (USFWS)

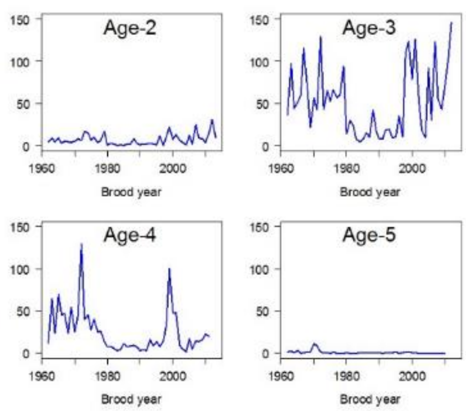
Spring Creek tule fall Chinook salmon



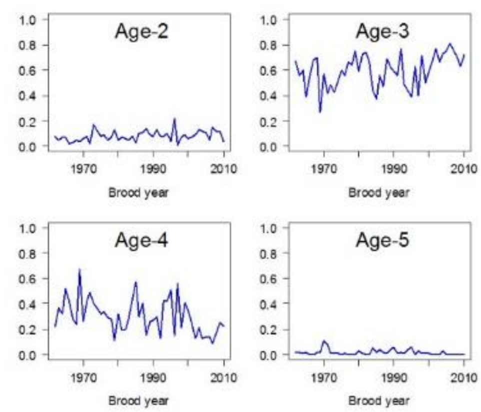
Spring Creek tule fall Chinook salmon forecast summary statistics

Forecast source and timeframe	Mean Return	MRE	MPE	MAE	MAPE
TAC 1980-2015	65,700	-2,000	7.8%	17,800	31.2%
TAC 1999-2013	99,100	-16,200	-7.3%	31,400	31.5%
CTC 1999-2013	99,100	-20,800	-13.9%	33,300	33.3%

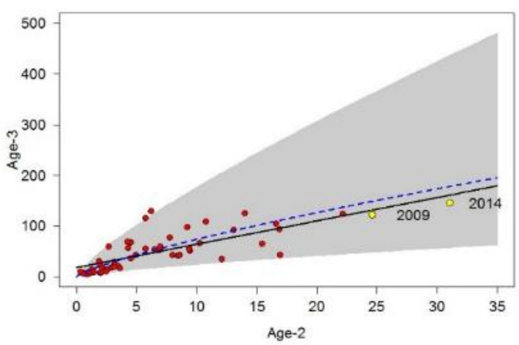
Spring Creek returns by age-class (BY 1962-2013)



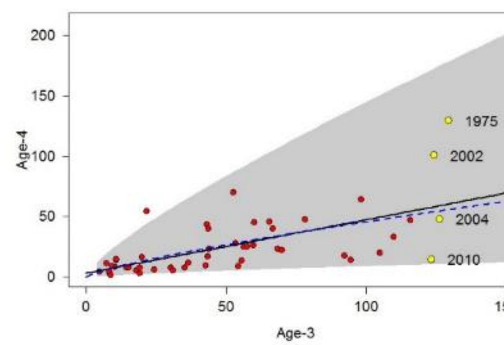
Spring Creek age-class proportions (BY 1962-2010)



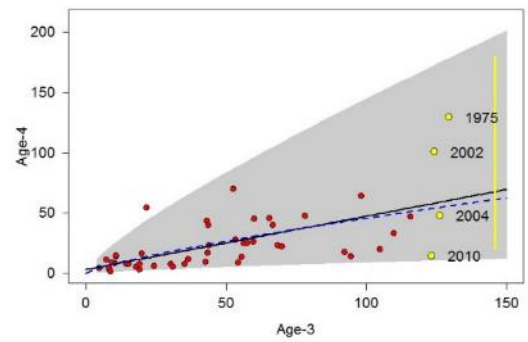
Spring Creek sibling relationships (BY 1962-2012)



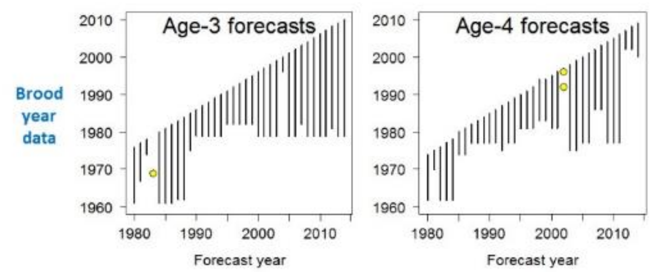
Spring Creek sibling relationships (BY 1962-2012)



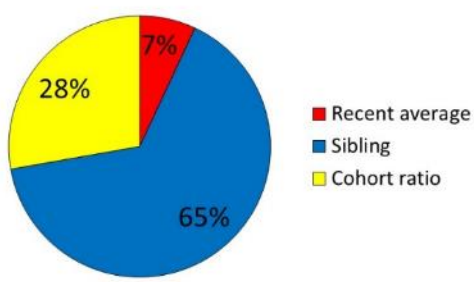
Spring Creek sibling relationships (BY 1962-2012)



Data used for Spring Creek forecasts



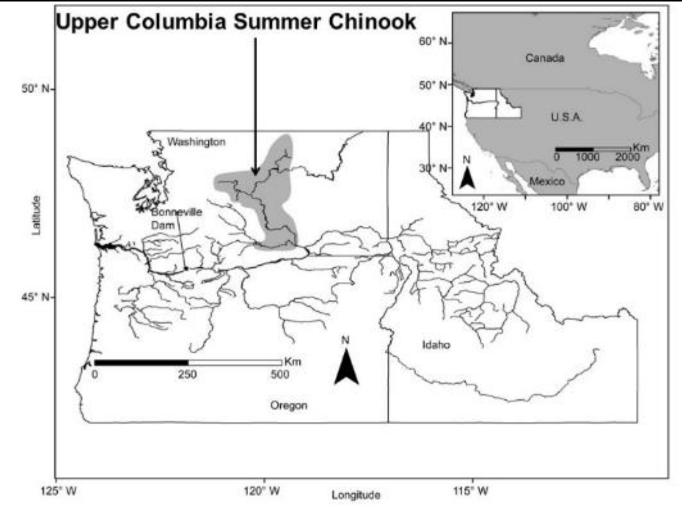
Models selected for forecasting age-3 and -4 returns (1980-2015)



Spring Creek Hatchery conclusions

- Natural variability in age-composition makes forecasting challenging
- High quality data on abundance and age-composition
- Forecasts have been unbiased, and demonstrated good precision

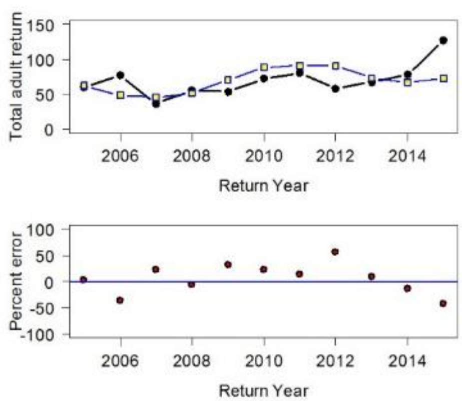
Columbia River Summer Run Forecast – Stuart Ellis (CRITFC)



Upper Columbia Summer Chinook

- Far-north migrating
- Mix of yearling and subyearling life histories; ocean-type and stream-type
- Mix of hatchery and wild
- Destined for areas upstream of Priest Rapids Dam and the Yakima River
- Recent 10-year average return to Columbia River is approximately 71,000 (range 37,000- 127,000)

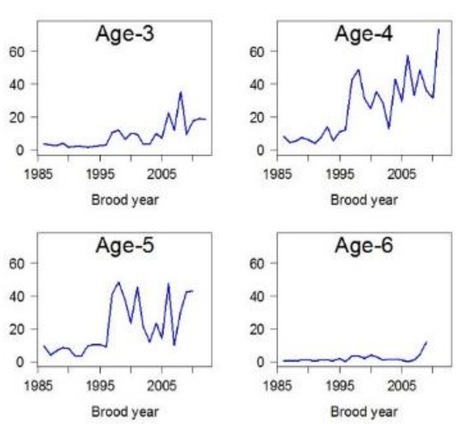
Upper Columbia summer Chinook salmon



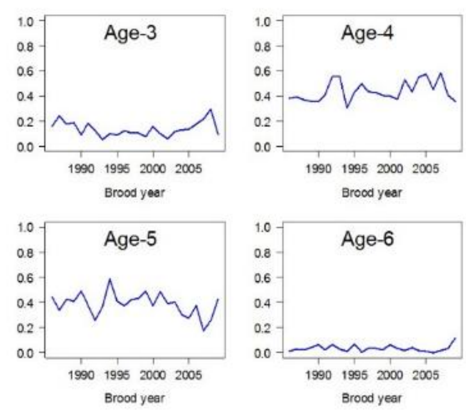
Upper Columbia summer Chinook salmon forecast summary statistics

Forecast source and timeframe	Mean Return	MRE	MPE	MAE	MAPE
TAC 2005-2015	69,800	-300	5.4%	17,300	23.5%
TAC 2005-2014, w/o 2006	62,600	8,800	15.4%	12,000	19.9%
CTC 2005-2014, w/o 2006	62,600	9,000	17.3%	12,100	21.8%

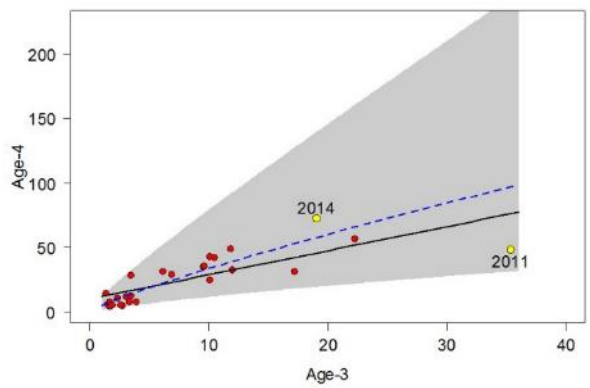
Summer Chinook returns by age-class (BY 1986-2012)



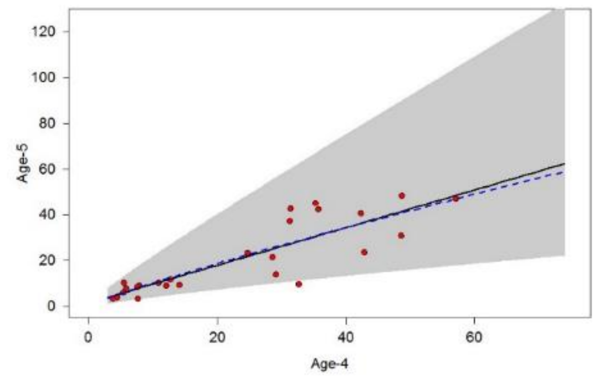
Summer Chinook age-class proportions (BY 1986-2009)



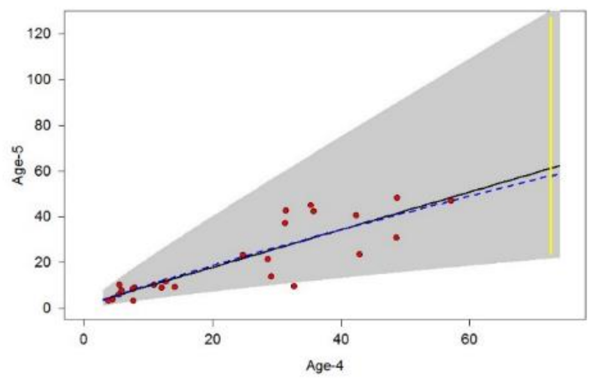
Summer Chinook sibling relationships (BY 1986-2012)



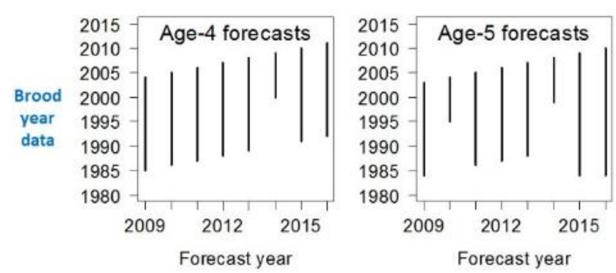
Summer Chinook sibling relationships (BY 1986-2012)



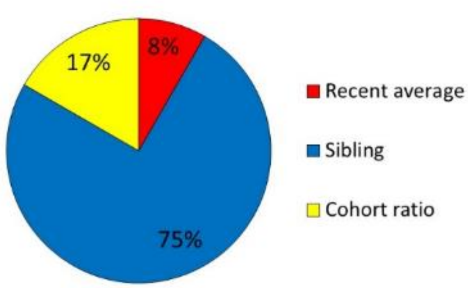
Summer Chinook sibling relationships (BY 1986-2012)



Data used for summer Chinook forecasts



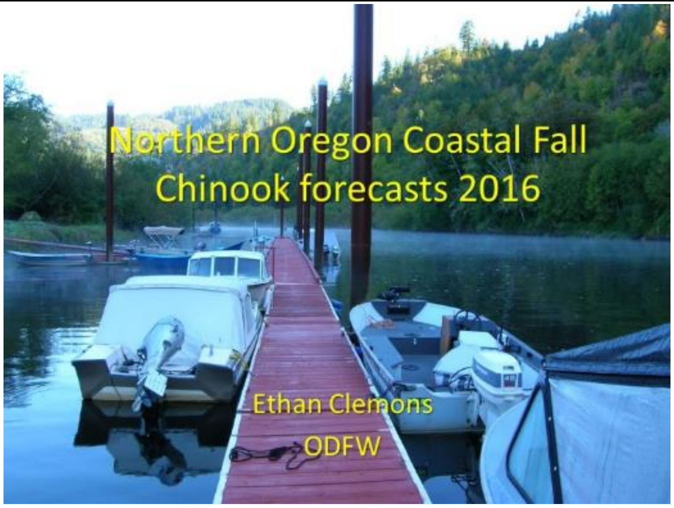
Models selected for forecasting age-3, -4, and -5 returns (2005-2016)

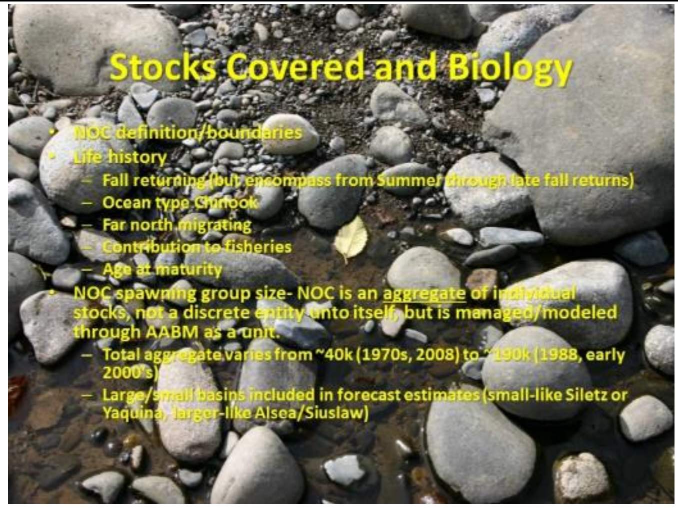


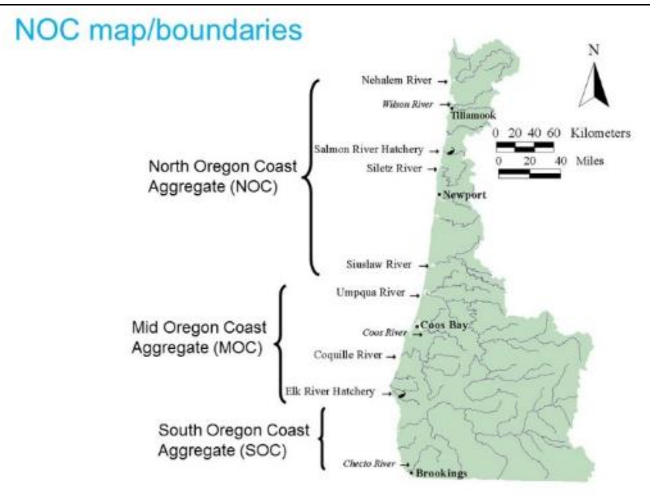
Summer Chinook conclusions

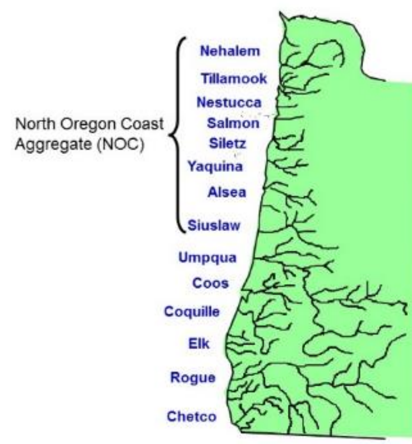
- Natural variability in age-composition makes forecasting challenging
- High quality data on abundance and age-composition
- Forecasts have been unbiased, and demonstrated good precision

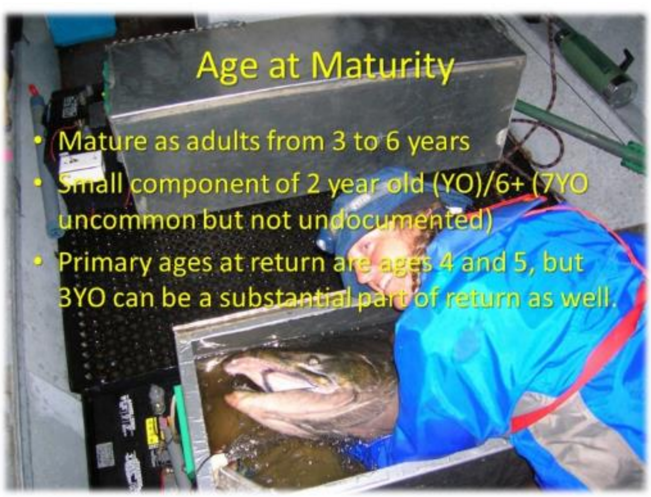
Northern Oregon Coast Fall Forecasts – Ethan Clemons (ODFW)

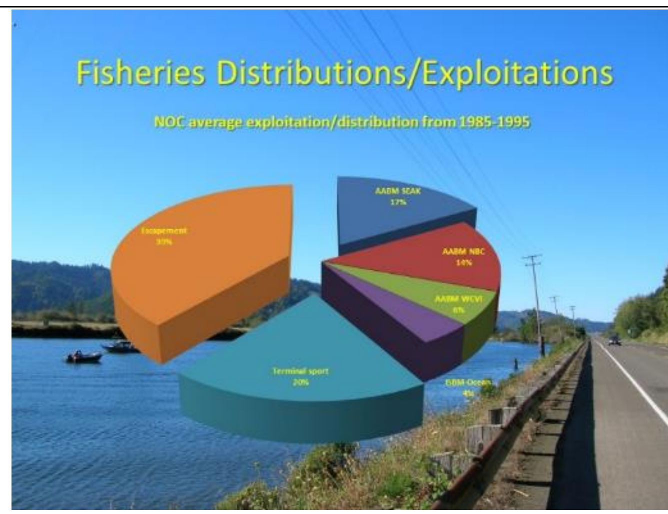


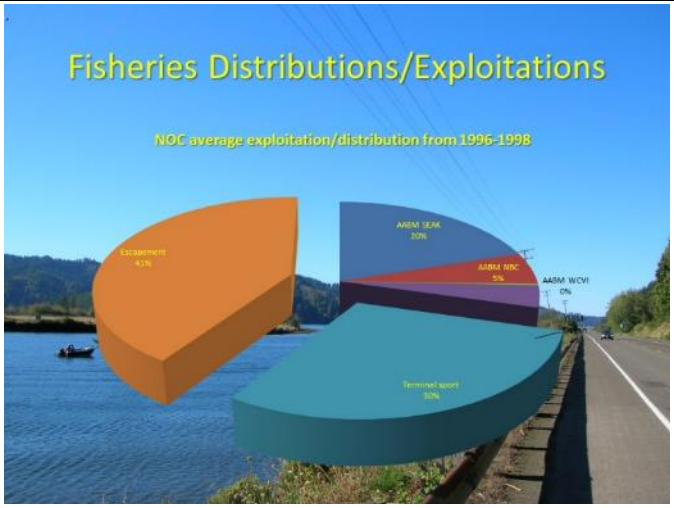


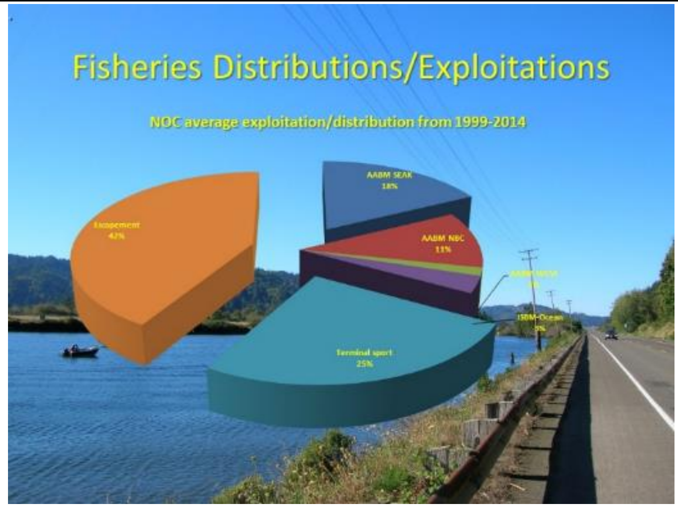




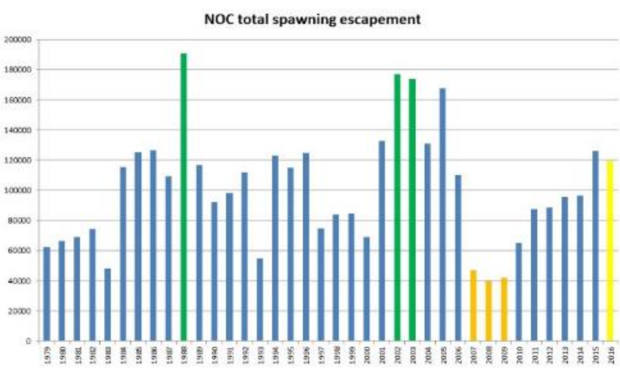








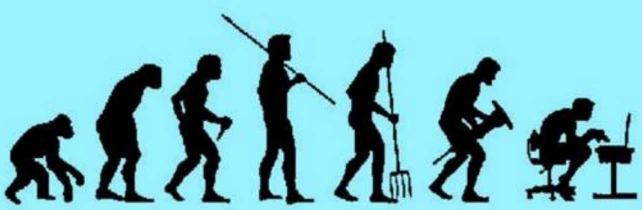
NOC Aggregate Size Through Time



Management Timelines

- Forecast due/used by mid-march
- Terminal returns of NOC spanning July-December
  - Scale age composition from field sampling
    - Creel surveys
    - Netting/weir/ M/R operations
    - Spawning ground surveys*
      - Predominant method of scale/age collection
    - Different programs, grants, supervision and goals from sampling programs
    - All field samples compiled/coordinated through Corvallis age reading lab
  - Sample size from NOC streams may vary from hundreds to tens of thousands between years.
    - Dependent predominantly on those grants secured through PSC for Chinook salmon enumeration/research projects
- Logistically, from sample to model projection; less than two months.
  - I usually get one to two weeks to generate forecasts before they are going to be used.

History of Forecasts for NOC



From Rudimentary to More Complex

- During 2008 renegotiation, ODFW was made an example of as a poor performer for stock forecasting…specifically for the NOC aggregate
  - 3 year average of escapement was being used as the method to derive forecasts for NOC aggregate
  - Poor forecast performance had been noted and documented for some time in CTC Calib & ERA reports (not a news-flash to most)
  - Co-incident with coastwide stock failures due to Oceanic impacts of the 2005 death brood.
- Prior to 2008, coastal fall Chinook forecasts were not an agency priority
  - Precipitous decline in escapement observed from 2007-10
  - Age sampling/reading which had occurred would not be reported for over a year after return had ended
  - No agency-wide coordination of age sampling programs/sampling projects
  - Age was an afterthought

From Rudimentary to More Complex-II

- 2008 saw changes to agency priorities and allowed for rapid-turnaround of scale age reads
  - Sibling regression –based forecasts which used recent observations of age composition to allow for more sophisticated modelling of projected returns began for NOC stocks in 2008.
  - Since 2008, each year's forecast performance has been monitored, analyzed, weighed, debated and (hopefully) improved.
    - Several different sib-regression relationships considered
    - Spurious conclusions routed-out
    - No one method has been consistently applied to all stocks in all years.
  - ForecastR
    - Modules came on-line and into use (past user testing) in 2016.
      - ARIMA models used in some cases for some stocks; in timeframe for PSC utilization
      - Post-season projections have continued using Naïve and Exponential smoothing models

Assumptions Made in Current Sibling Regression Forecasts

- Current models are naïve to fisheries' impacts
  - Starting with the statement "assuming all fisheries are going to have about the same proportional impact as they have had during the time series used to generate these relationships..."
  - Inability to forecast the impact of Northern Fisheries (or terminal fisheries for that matter) in usable timeframe.
- Assume maturation will be comparable to observed values
- Assume year to year survival (0-2, 2-3, 3-4...) will be static
- Assume age specific sampling is non-biased (or correct for known biases).
- Assume projections will be within the bounds of previously observed values.
- Forecast for escapement, not terminal return
  - Explanation as to why
- Assume static relationship between surveyed streams and unsurveyed streams/basins (expansion by 17%).

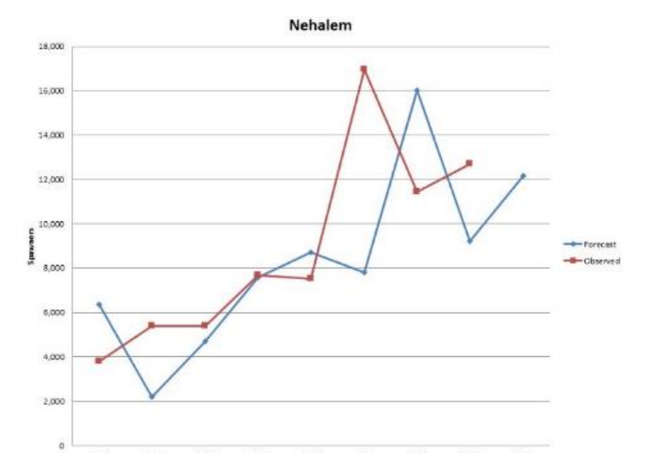
How are NOC Forecasts Currently Calculated?

- Individual basins are forecasted
  - Naïve sibling regressions cast through timeseries
  - Additional expansions for unsurveyed areas
- Aggregated forecast for NOC
  - All 7 basins added together
  - Compared to aggregated forecast
- Recent additional methods available through Forecast R
  - Exp smoothing, Naïve, ARIMA (Mechanistic, Return Rate, Simple Sib Regressions, Log-power Sib regressions, Complex Sib Regressions...all waiting in the wings through Phase 3), plus the potential for covariate incorporation.
  - Not only end products, but readily available calculation & documentation of performance.

2016 forecasts for NOC

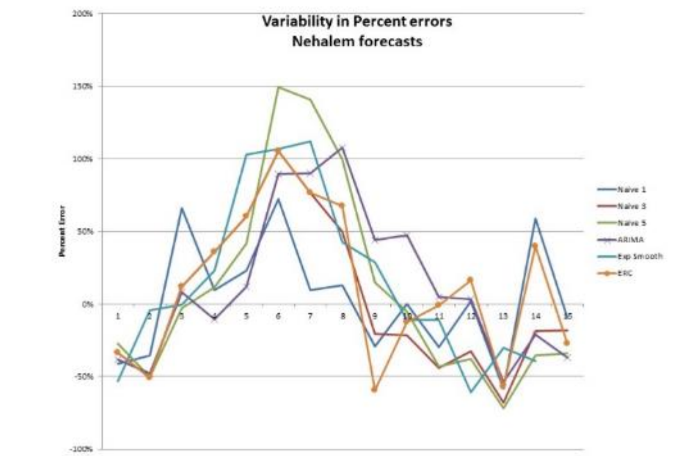
- Mixture of both traditional sib-regressions and ARIMA generated forecasts

Run year	Nehalem	Tillamook	Nestucca	Siletz	Yaquina	Alsea	Siuslaw	total
2016 forecast sibling ERC	12,165	16,109	9,051	7,528	15,813	45,531	40,840	147,038
2016 forecast ARIMA	9,629	12,673	8,481	5,362	6,992	27,259	28,701	99,097
							hybrid	119,374



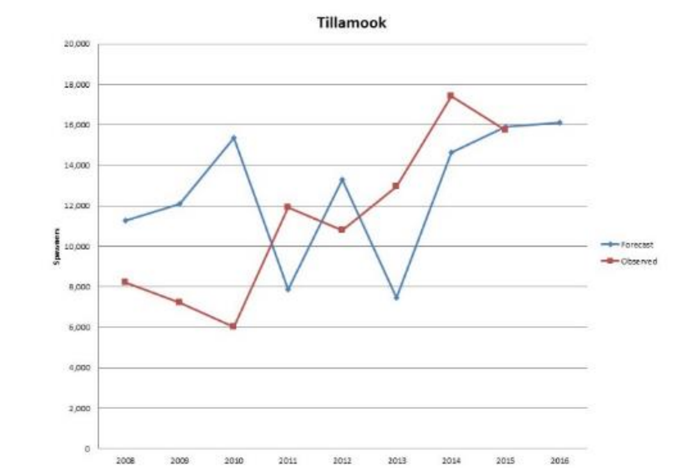
Nehalem Forecast Diagnostics

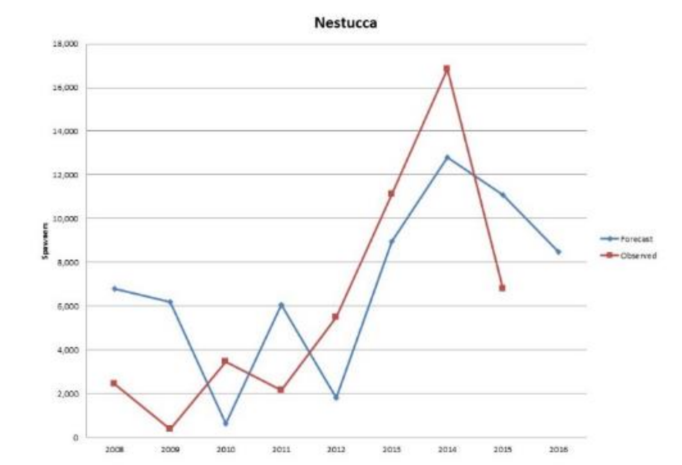
Return Year	Nehalem esc Observed values	Naïve 1 Forecast	Naïve 1 Dev	Naïve 1 Rel Error (%)	Naïve 3 Forecast	Naïve 3 Dev	Naïve 3 Rel Error (%)	Naïve 5 Forecast	Naïve 5 Dev	Naïve 5 Rel Error (%)	ARIMA Forecast	ARIMA Dev	ARIMA Rel Error (%)	Exp Smooth Forecast	Exp Smooth Dev	Exp Smooth Rel Error (%)	ERC Forecast	ERC Dev	ERC Rel Error (%)
2001	11,662	6,856	-4,806	-41%	7,722	-3,940	-34%	8,480	-3,182	-27%	7,165	-4,497	-39%	7,775	-3,887	-33%	7,722	-3,940	-34%
2002	18,089	11,662	-6,427	-36%	8,860	-9,229	-51%	8,970	-9,119	-50%	9,443	-8,646	-48%	8,494	-9,595	-53%	8,860	-9,229	-51%
2003	10,906	18,089	7,183	66%	12,202	1,296	12%	10,583	-323	-3%	11,807	901	8%	10,436	-470	-4%	12,202	1,296	12%
2004	9,976	10,906	931	9%	15,552	5,577	56%	11,115	1,140	11%	8,914	-1,061	-11%	9,925	-52	-1%	15,552	5,577	56%
2005	8,114	9,976	1,861	23%	12,990	4,876	60%	11,497	3,383	42%	9,088	974	12%	9,973	1,859	23%	12,990	4,876	60%
2006	4,711	8,114	3,403	72%	9,665	4,954	105%	11,749	7,038	149%	8,920	4,209	89%	9,569	4,858	103%	9,665	4,954	105%
2007	4,304	4,711	407	9%	7,602	3,298	77%	10,360	6,056	141%	8,185	3,881	90%	8,900	4,596	107%	7,602	3,298	77%
2008	3,810	4,304	494	13%	5,713	1,903	50%	7,604	3,794	100%	7,911	4,101	108%	8,071	4,261	112%	6,378	2,569	67%
2009	5,390	3,810	-1,581	-29%	4,278	-1,113	-21%	6,184	794	15%	7,770	2,380	44%	7,693	2,303	43%	2,186	-3,204	-59%
2010	5,384	5,390	6	0%	4,226	-1,158	-22%	5,102	-282	-5%	7,925	2,541	47%	6,916	1,532	28%	4,708	-676	-13%
2011	7,666	5,384	-2,281	-30%	4,270	-3,395	-44%	4,367	-3,298	-43%	8,037	372	5%	6,811	-854	-11%	7,578	-87	-1%
2012	7,816	7,666	150	2%	5,057	-2,458	-33%	4,658	-2,857	-38%	7,762	347	3%	6,684	-831	-11%	8,728	1,212	16%
2013	18,194	7,816	-10,678	-59%	5,795	-12,399	-68%	5,151	-13,043	-72%	8,388	-9,806	-54%	7,131	-11,063	-61%	7,815	-10,378	-57%
2014	11,482	18,194	6,741	59%	9,328	-2,134	-19%	7,397	-4,055	-35%	9,043	-2,419	-21%	7,996	-3,456	-30%	15,998	4,546	40%
2015	12,678	11,482	-1,226	-10%	10,391	-2,287	-18%	8,357	-4,321	-34%	8,009	-4,669	-37%	7,691	-4,987	-39%	9,225	-3,453	-27%
	CumulErr		-6,822			-18,198			-18,278			-11,492			-16,788			-4,638	
	Mean Error (ME)		-388			-1,213			-1,218			-766			-1,082			-309	
	Mean Relative Error (MRE)			3%			2%			10%			13%			11%			11%
	MAPE			31%			43%			81%			41%			46%			44%
	RMSE		4,480			4,910			5,365			4,348			4,778			4,715	
	MAE		3,212			3,867			4,179			3,380			3,640			3,820	
2016		12,678			12,380			11,501			9,629			8,702			12,165		

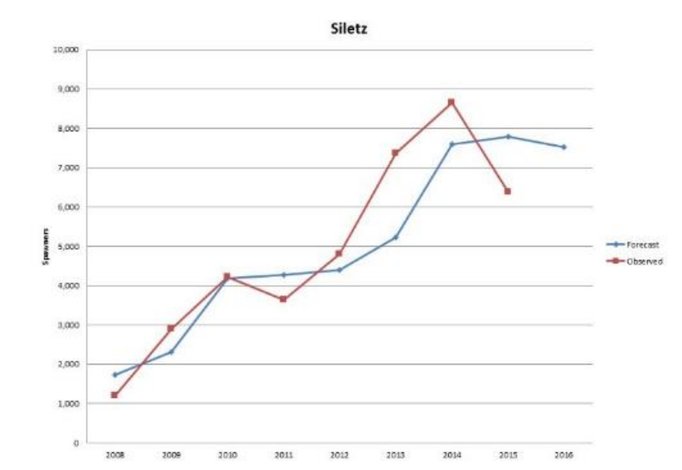


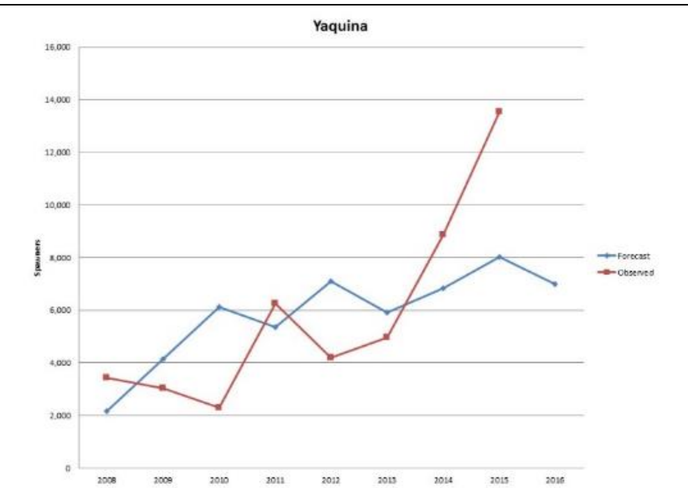
ForecastR Diagnostics

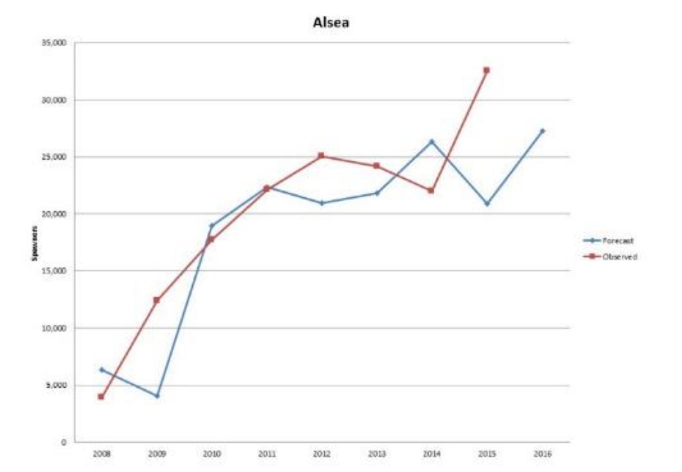
- 5 different models
- 7 different stocks
- 4 different age classes
- 1500pp (+) of documentation available for those stocks comprising NOC
  - Available upon request
- Documentation on current forecasting methods (Labeled "ERC") sparse to nonexistent
  - Done in extraordinarily compressed time scale(one to two week turnaround)
  - Previous comparison between ForecastR sibling regression and ERC very similar results

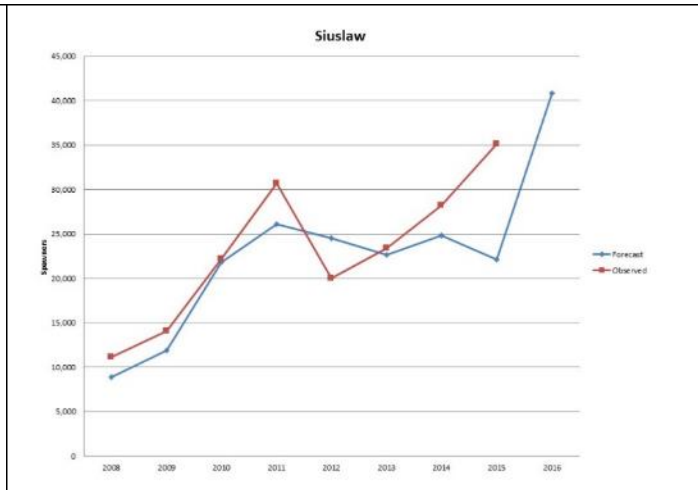


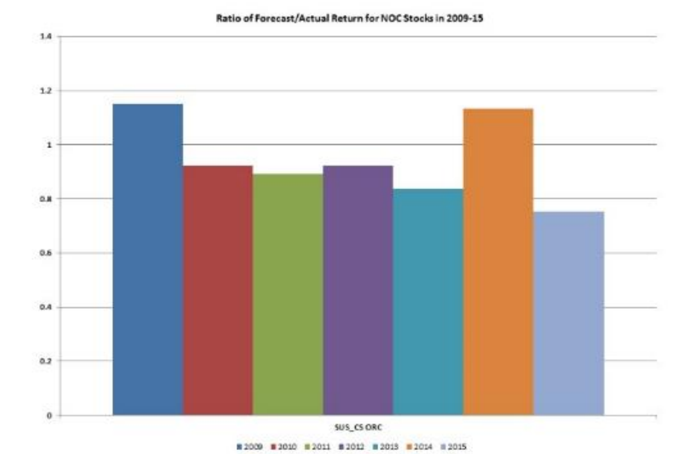












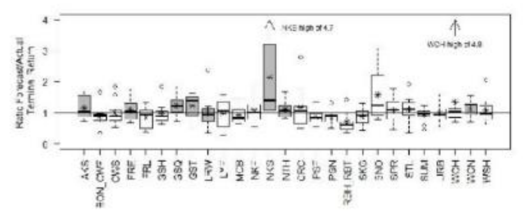
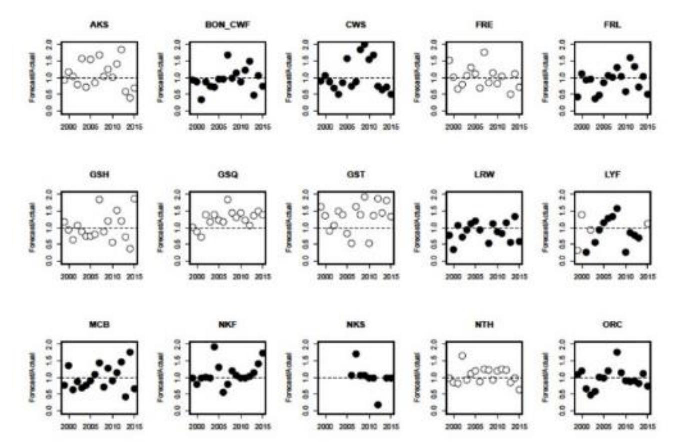
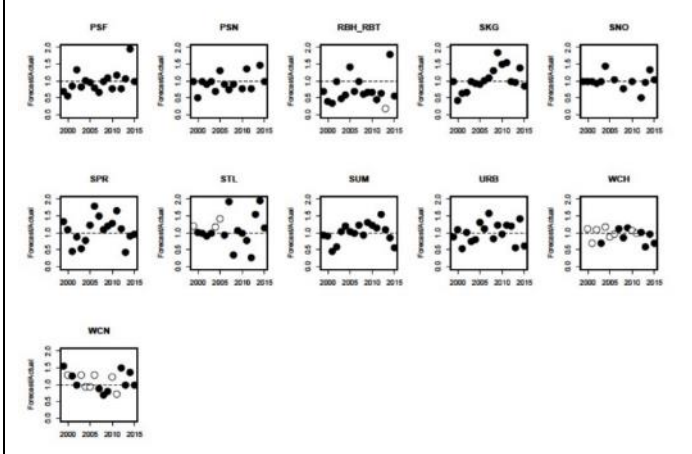


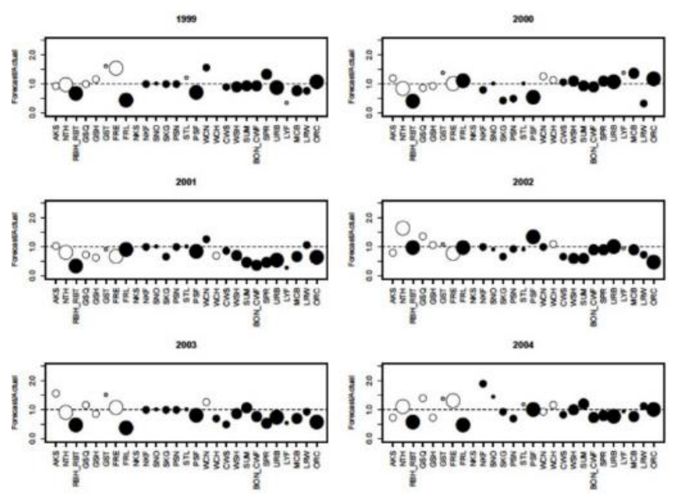
Figure 13—The ratio of the forecasted terminal return to actual terminal return from 1999–2008 for stocks used in the CTC model calibration.

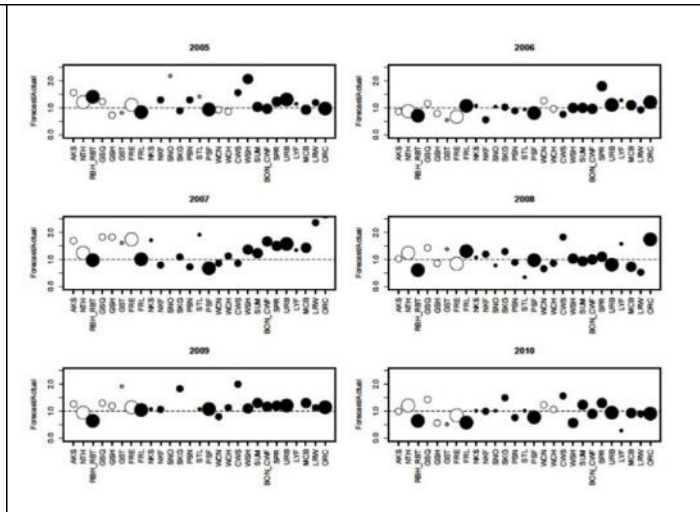
*Shaded boxes are for stocks that have no agency supplied forecasts. Stock abbreviations are in Table 14. The median error bar is the middle of the box. 50% of the errors are above the line, 50% below. Lower and upper ends of the boxes are the 25th and 75th percentiles, respectively. Mean ratios are represented as *, top and lower vertical bars are values that are 1.5 greater (or less than) the values at the end of the boxes. Open circles (O) represent outliers, values that are beyond upper and lower bars.*

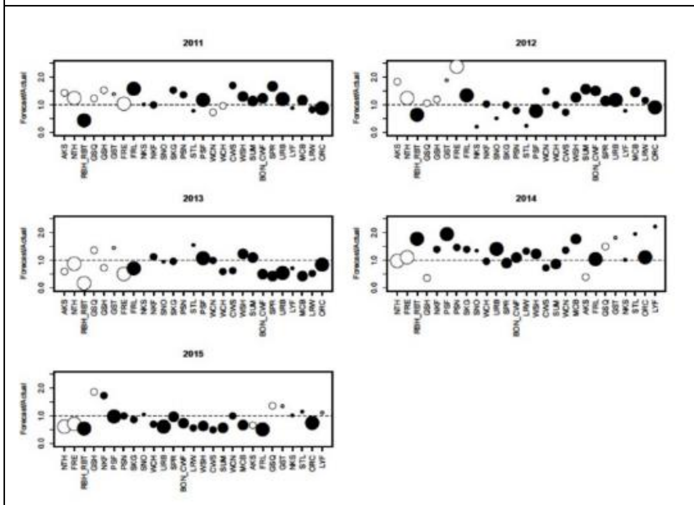
Page 58











Next Steps

- Correction of non-surveyed areas expansion
  - As long as we're using the same base period/habitat expansion methods to derive escapement
- Evolution of escapement estimation methods to M/R based and Peak Count Model
  - Whole time series will be re-worked
  - Concurrent effort to update age composition/escapement at age both historically and prospectively
- Utilization of ForecastR models (addition of complex sibling, other potential models to add to the arsenal, external co-variates (age 2 survival, etc.) into forecast projections).

Reasonable Expectations

- Forecasts will be wrong; we're not going to hit the target all the time.
- Our aim is to miss less and less with each passing year
- Utilize the available information to the best possible outcome
  - No matter how sophisticated our modelling system, those products are still going to rely on:
    - Accurate spawning escapement estimates
    - Accurate age composition estimates
    - Both will be direct products of investments into field surveys

Potential Criteria for Forecasts <u>Both</u> Agency and Chinook Model Based

- Consistent pattern of bias observed for X+ years (3 years?)
- What constitutes "better"?
  - 10%, 15% better?, MAPE?, RMSE? accuracy?
  - Volatility
- <u>Bi-directional</u>, not just higher or lower yielding forecasts.
- Time period chosen for retrospective evaluation will have acute bearing on conclusions
  - Change in methods, data through time

Observations to Take Into Next Year

- Virtual non-presence of jacks in 2015 NOC and MOC return
  - A total of 6 jacks were observed coast-wide (both NOC and MOC) for RY 2015
  - Doesn't bode well for 3 YO return in 2016 and following cohorts
- Blob/El Nino impacts looming on horizon
- Current models do not account for this external knowledge/observations

Closing Slide: Good News for the Future!?!

- Grant money for field projects in NOC has taken precipitous decline in recent years
  - Field sampling has dropped to minimal/non-existent in some basins
- ODFW Aging lab had recently weathered retirement/personnel turnover in past years
  - Lab has lost a FTE
- We have less money, fewer samples and fewer people to derive estimates that will be facing increasing scrutiny with greater economic impact on an increasingly accelerated time scale.

ForecastR Statistical Package - Antonio Velez-Espino (CDFO)

ForecastR (Phase-3)

ForecastR Motivation

- Expediting completion of forecasts of abundance of salmon stocks (originally for Chinook only)
- Produce a unified forecasting tool which can be used by researchers and managers across different jurisdictions
- Facilitate communication and sharing of forecasting results (transparency)
- Incorporate a variety of statistical modeling and forecasting tools with the aim to improve quality of forecasts

ForecastR Mechanics

- ForecastR relies on the open source **R** software available at www.r-project.org
- ForecastR will provide a graphical user interface (**GUI**) to facilitate interaction with the users
- Users will not need to know how to code in R to use **ForecastR**
- ForecastR's GUI is being built with functionality available via the R package **PBSmodelling**

ForecastR Features

- **ForecastR** allows users to forecast abundance of individual stocks (e.g., Chinook, Chum, Coho) based on historic data and other available information
- **ForecastR** primarily accommodates two types of time series (age-specific or total abundance) representing individual stocks or aggregates

ForecastR Features

- ForecastR recognizes three different kinds of abundance metrics:

Escapement (ESC)

Terminal Run (TR)

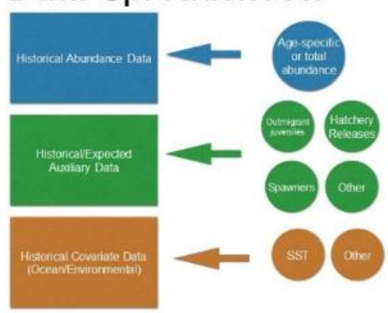
Production (PROD)

ForecastR Workflow

Step 1:	Enter Input Information
Step 2:	Conduct Statistical Modeling
Step 3:	Check Modeling Diagnostics
Step 4:	Produce Abundance Forecast(s)
Step 5:	Assess Retrospective Forecast Performance
Step 6:	Check Forecast Diagnostics
Step 7:	Report Results

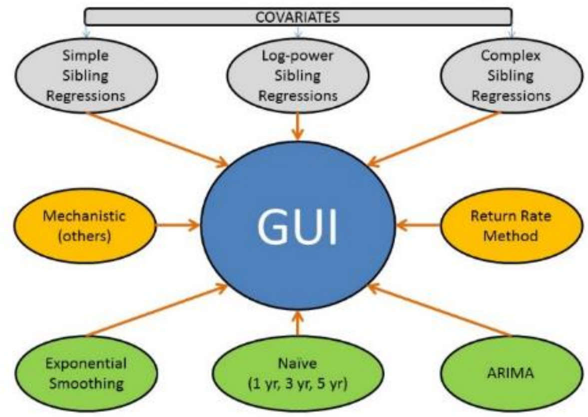
Step 1: Enter Input Information

Data Spreadsheets



7

Step 2: Conduct Statistical Modeling



8

Step 3: Check Modeling Diagnostics

- Non-linearity of the response-predictor relationship(s)
- Correlation of the model errors
- Non-constant variance of the error terms
- Non-normality of error terms
- Outliers
- High-leverage points
- Collinearity
- Time series plot of the model residuals
- Autocorrelation plot of the model residuals
- Partial autocorrelation plot of the model residuals
- Plot of p-values associated with the Ljung-Box test applied to the model residuals

9

Step 4: Produce Abundance Forecasts

Given
- Relevant *input information*
- An *individual stock*
- A specified *abundance metric*
- A chosen *model*

ForecastR produces:

Terminal Run	Best Model	Forecasting Year	Point Forecast	Interval Forecast
Age_2	ARIMA(0,0,0) with non-zero mean	2013	6,449	1,000 - 15,400
Age 3	Age_3 ~ -1 + Age_2	2013	25,982	17,779 - 68,025
Age 4	Age_4 ~ -1 + Age_3	2013	20,932	8,373 - 42,129
Age 5	Age_5 ~ -1 + Age_4	2013	429	170 - 1,719
Total	-	2013	53,792	40,826 - 110,650

10

Step 5: Assess Forecast Performance

ForecastR computes 6 retrospective forecast performance measures:

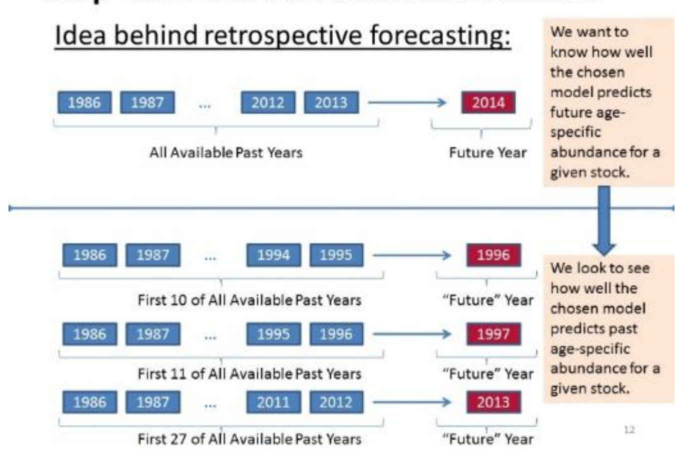
ID	Measure	Name
1	MRE	Mean Raw Error
2	MAE	Mean Absolute Error
3	MPE	Mean Percent Error
4	MAPE	Mean Absolute Percent Error
5	RMSE	Root Mean Square Error
6	MASE	Mean Absolute Scaled Error

Used to characterize the *central tendency* and *variability* in the distribution of annual retrospective forecast errors

11

Step 5: Assess Forecast Performance

Idea behind retrospective forecasting:



12

Step 5 (cont'd): Ranking of models for getting "best" point forecast

Given an abundance stock and an abundance metric, find the "best" model for each age class among a set of competing models for producing point forecasts of stock abundance. ("Best" model yields "best" point forecast.)

Age k	MRE	MAE	MPE	MAPE	RMSE	Average Rank
Model 1	1	2	1	4	3	2.2
Model 2	5					
Model 3	4					
Model 4	3					
Model 5	4					

Step 1: Rank all models with respect to the values of each performance metric (e.g., MRE).

Step 2: Compute the average rank corresponding to each model.

Note: User can choose which metrics are meaningful to them (e.g., MRE, MPE and RMSE only) and average ranks only across those metrics.

Step 5 (cont'd): Ranking of models for getting "best" point forecast

After applying steps 1 and 2 described on previous slide for all available age classes, we'll end up with:

Model	Average Model Rank Across Performance Metrics for Each Age Class			
	Age 2	Age 3	Age 4	Age 5
Model 1	2.2			
Model 2	1.0			
Model 3	1.5			
Model 4	3.0			
Model 5	3.1			

Step 3: For each age class, identify the model with the **lowest** average ranking as the "best" model for the purpose of point forecasting.

Note: If two or more competing models have the same average rank, we can use "model averaging".

Step 5 (cont'd): Ranking of models and total abundance

Approach No. 1

a) Find "best" model for Age 2
b) Find "best" model for Age 3
c) Find "best" model for Age 4
d) Find "best" model for Age 5
e) Compute total abundance by adding up the point forecasts produced by the "best" model for each age

Approach No. 2

a) Compute total abundance by adding up the point forecasts produced by the same model

Model 1:
Age 2 + Age 3 + Age 4 + Age 5 = **Total [1]**
Model 2:
Age 2 + Age 3 + Age 4 + Age 5 = **Total [2]**
.
.
.
Model 5:
Age 2 + Age 3 + Age 4 + Age 5 = **Total [5]**

b) Rank models to find the "best" point forecast of total abundance across models.

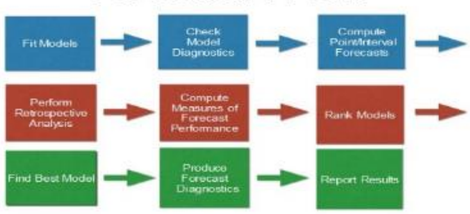
Step 6: Perform Forecast Diagnostics

ForecastR includes visual and numeric forecast diagnostics, such as:

- ✓ Histograms/boxplots of retrospective forecast errors
- ✓ Scatterplots of forecasted versus actual stock abundance values
- ✓ Percentage of retrospective forecast intervals that capture the actual stock abundance value (showing how well models are doing over time)
- ✓ Probability Profiles

Step 7: Report Results

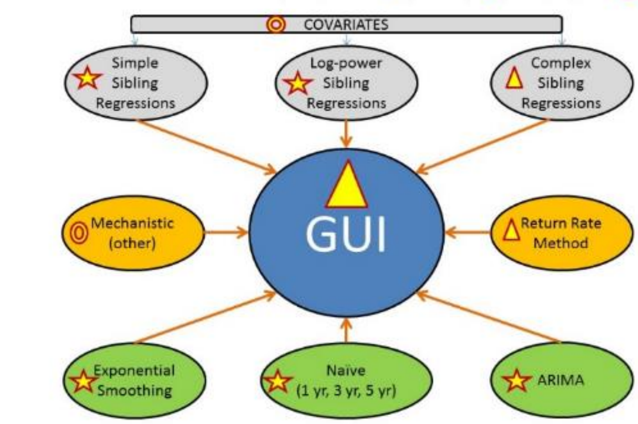
ForecastR Flow



ForecastR produces two types of reports:

A. Executive Summary
B. Full Report

INDIVIDUAL-MODULES PROGRESS (completed: ★ ; partially: Δ ; little progress: ◎)

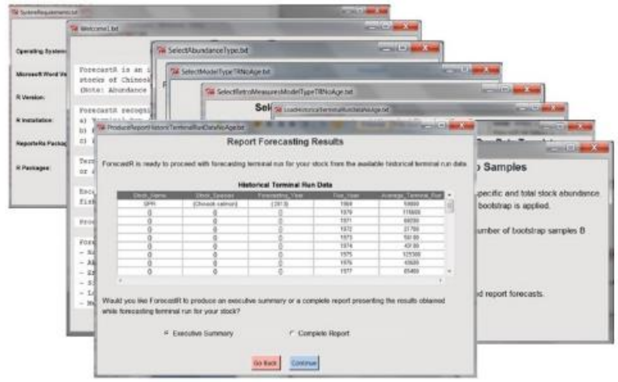


Individual-Modules capabilities

- Produce Word or Html reports (include table of contents, numbered figures and tables with captions, and stats tutorials)
- Point forecast and bootstrapped-based interval forecast
- Numerous diagnostics
- Model selection (ARIMA, Exponential Smoothing, and Complex Sibling Regressions)
- Probability profiles
- Retrospective evaluations of model performance
- Model Ranking takes place externally

19

Appendix: Prototype GUI



20

Discussion of Agencies' Forecasts – Facilitated Discussion

Table 5. Central issues or improvements identified by workshop participants for consideration by the panel with respect to stock forecasts.

1.	Recent model performance (recent averages or weighted averages)
2.	Kalman filter sibling model
3.	Model selection criteria & definitions
4.	Be open to many alternative models (don't fall in love with a single model)
5.	More in-depth model review process (rigor and structure)
6.	Accommodate timeliness or lack thereof in the availability of information for use in forecasts
7.	Need CTC guidance on forecast requirements & definitions
8.	Rely on the data available (given limitations of existing resources)
9.	Fundamental data needs
10.	Practical guidance vs. generalities
11.	Consistency of fish currencies (forecast vs observed)
12.	Exploration of external information (environmental, maturation, size, juveniles, etc.)
13.	Empirical dynamic modeling
14.	Probabilistic framework
15.	Tradeoffs of accuracy vs. overfitting
16.	Apply this same kind of rigor to the PSC Chinook model (for the abundance index)

Panel Impressions – Randall Peterman

Wrap-up Comments

Pacific Salmon Commission's "Workshop on Chinook Forecasting", Portland, Oregon, 10-11 August 2016

1

Objective: Improve forecasts

Your suggestions + ours:

-One of the Panel's issues is the discrepancy between CTC's available resources and the Treaty's requirements (and short time constraints)

-Constrained by Table 1 of Treaty – possibly change in the current negotiations?

-Better documentation (CTC model + regional forecasts fed into that model)

2

Uncertainties in hypothesis/assumptions

- Structural uncertainty
- Parametric uncertainty
- Unclear management objectives
- Outcome uncertainty

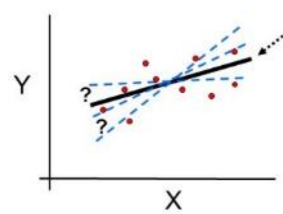
Ideas apply to CTC model and stock-specific forecasts

"Best practices"

3

Structural uncertainty - 2

1. Exploring alternative models reflects uncertainty

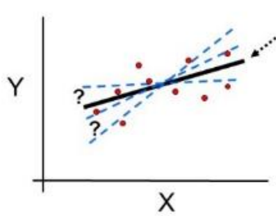


Best fit
If use only best-fit model implies:
Prob.(best-fit model is true) = 1.0
and
Prob.(all other models) = 0

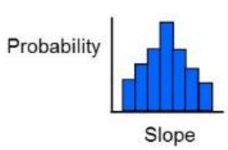
4

Structural uncertainty - 3

1. Exploring alternative models reflects uncertainty



Best fit
If use only best-fit model implies:
Prob.(best-fit model is true) = 1.0
and
Prob.(all other models) = 0



5

Structural uncertainty - 4

1. Explore alternative models
   - Sibling
     - -- Fit untransformed abundances
     - -- Fit $\log_e$(abundances)
     - -- Hybrid sibling (standard sibling + naïve) (Haeseker et al. 2007)

6

Hybrid sibling model (standard sibling + naïve)

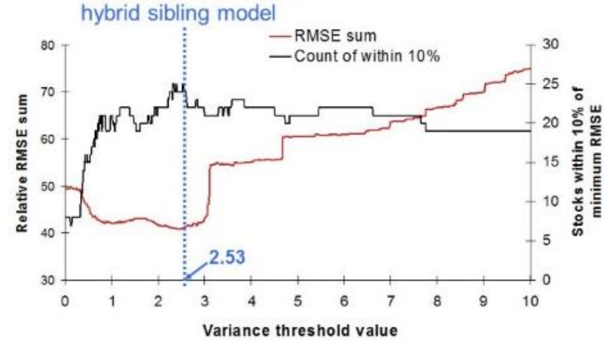
- Standard sibling models good when residual variance is LOW

but ...

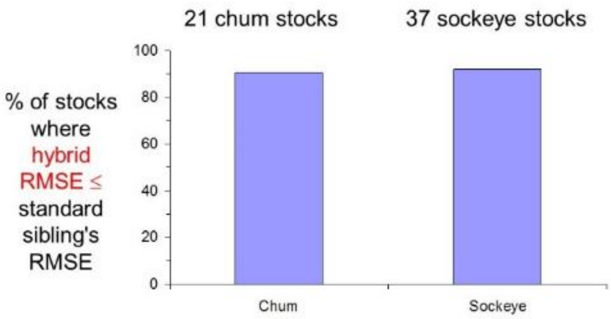
- Naïve models such as $R_t = R_{t-4}$ good when residual variance is HIGH
- Hybrid model:
  1. Uses standard sibling model if its residual variance < X
  2. Uses naïve model if variance $\geq$ X
- Search for optimum X

(Haeseker et al. 2007, NAJFM 27:634) 7

Optimal variance threshold for sockeye hybrid sibling model



(Haeseker et al. 2007, NAJFM 27:634) 8



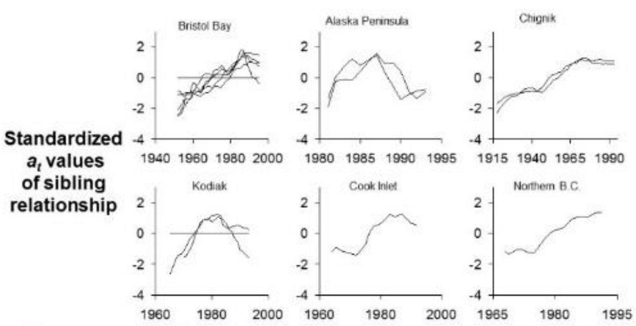
(Haeseker et al. 2007, NAJFM 27:634) 9

Structural uncertainty - 5

1. Explore alternative models
   - Sibling
     - -- Fit untransformed abundances
     - -- Fit $\log_e$(abundances)
     - -- Hybrid sibling (standard sibling + naïve) (Haeseker et al. 2007)
     - -- Estimate time-varying parameters: Kalman filter

10

Sockeye

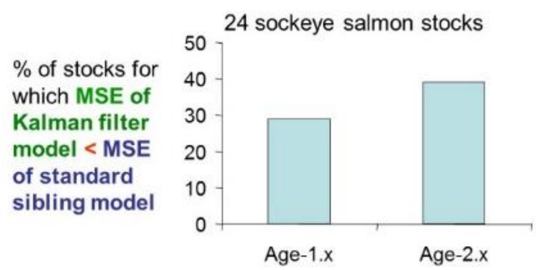


Brood year

Widespread shared trend

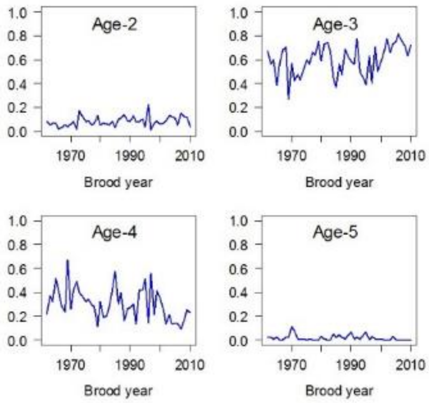
(Holt and Peterman 2004, CJFAS 61:2455) 11

Results of retrospective analyses



(Holt and Peterman 2004, CJFAS 61:2455) 12

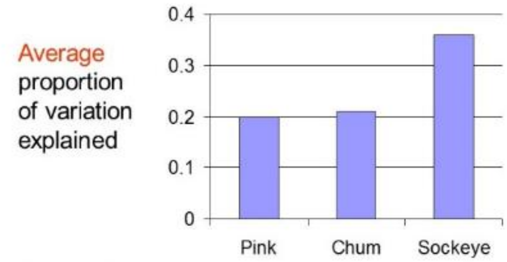
Good candidate: Spring Creek tule fall Chinook



(Whisler 2016, submission to this panel)

13

Using the best pre-season forecasting model for each stock



A lot of unexplained variation!

(Haeseker et al. 2005, 2008, NAJFM)

14

Structural uncertainty - 6

Explore alternative components of **CTC model**
- Effects of:
  - -- Body size on age at maturity
  - -- Ocean conditions on survival rate
  - -- Fleet's changing q, spatial distribution, ...
  - -- Covariation among stocks in matur. rate, surv. rate

15

Structural uncertainty - 6

Long term: develop alternative models such as
- 3 options: improve existing, substitute more complicated alternative, substitute simpler alternative that captures essential elements
- Continuous catch equations (e.g. Morishima & Chen)
- Statistical catch-at-age models (e.g. Sharma & Yuen)
- Management-strategy evaluation models (MSE) used in several marine fisheries and Fraser sockeye

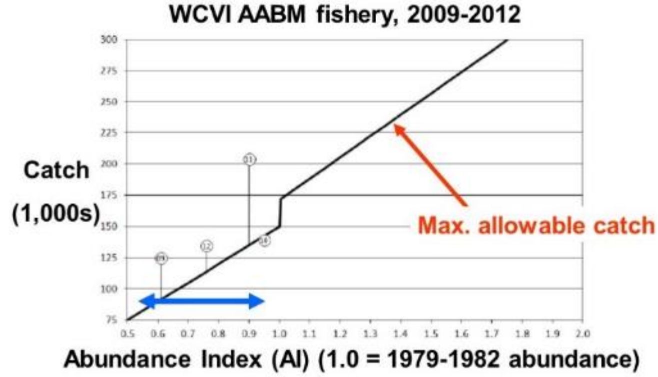
16

Structural uncertainty - 7

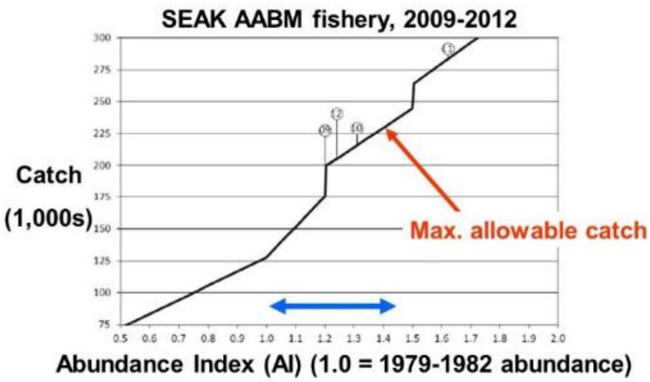
2. Retain forecasts from the alternative models
   - Prediction intervals on forecasts (CTC + regional)
     - -- Multi-model averaging based on AIC_c weights (Burnham and Anderson 2002)
     - -- Stock-specific forecasts and CTC's AABM AIs

17

WCVI AABM fishery, 2009-2012



p. 98 of TCCHINOOK 14-1_Vol1 (2014)



p. 97 of TCCHINOOK 14-1_Vol1 (2014)

Structural uncertainty - 8

ForecastR – Excellent!
Extensions:
- Hybrid sibling model
- Kalman filter estimation of time-varying sibling parameters
- Multi-model averaging based on AIC_c weights

20

Parametric uncertainty - 2

- Imperfect data from observation error
- Non-stationarity (changing mean and/or variance)
- Variations over time and/or space in:
  - -- Maturation rate
  - -- Survival rate / productivity
  - -- Exploitation rates
  - -- Fishing Policies (FPs), efficiency of fishing gear
  - -- Proportion of fish vulnerable

21

Parametric uncertainty - 2

What can we do?
- Better data
- Use covariation among stocks in ...?
- Data Generation Model (DGM) – Excellent!
- Scenarios of ranges of parameters (ensembles)

22

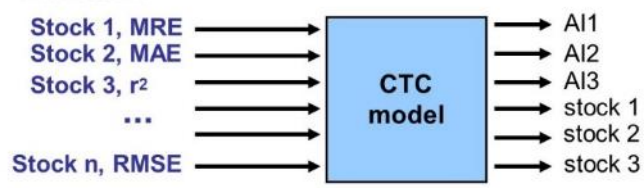
Unclear management objectives

Rank order of alternative forecasting models can be affected by ranking criteria:
- Minimize bias
- Maximize precision
- Minimize prob. of large errors (e.g. >30%)

Need consistency across years within a region and across regions.

23

Stock-specific forecasts based on model with best ...



24

Unclear management objectives

Alternative stock-specific forecasting models should be ranked using criteria that match managers' objectives.

How can managers use information on uncertainties in AIs?

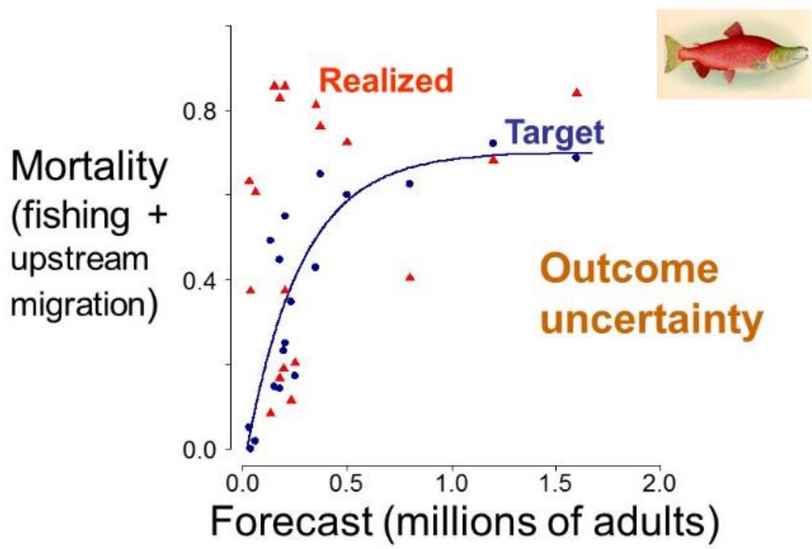
25

Outcome uncertainty

- Deviations from targets
(e.g., desired vs. realized exploitation rates)

26

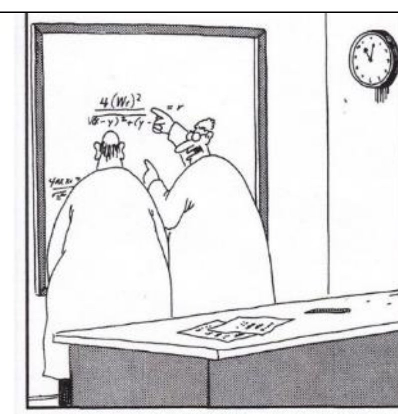
Early Stuart sockeye salmon, B.C. (1986-2003)



(Holt and Peterman 2006)

If all else fails …

28



"Yes, yes, I *know* that, Sidney…*everybody* knows *that!*…But look: Four wrongs *squared*, minus two wrongs to the fourth power, divided by this formula, *do* make a right."

29

Appendix E - List of Acronyms & Abbreviations

AABM	Aggregate Abundance Based Management
ADF&G	Alaska Department of Fish and Game
AEQ	Adult Equivalent
Agreement	June 30, 1999 PST Annex and the related agreement
AI	Abundance Index
AIC	Akaike Information Criteria
APC	Average Proportion Correction procedure
AUC	Area-Under-the-Curve
AWG	Analytical Working Group of the CTC
BC	British Columbia
BY	Brood Year
BYER	Brood Year Exploitation Rate
CBC	Central British Columbia (Kitimat to Cape Caution)
CDFO	Canadian Department of Fisheries and Oceans
CI	Confidence Interval
CLB	Calibration
CNR	Chinook Nonretention
CPUE	Catch per unit effort
CR	Chinook Retention
CRITFC	Columbia River Inter-Tribal Fisheries Commission
CTC	Chinook Technical Committee
CV	Coefficient of Variation
CWT	Coded Wire Tag
CWTIP	Coded Wire Tag Improvement Program
CWTIT	Coded Wire Tag Improvement Team
CY	Calendar Year
CY	Catch Year
DFO	Department of Fisheries and Oceans Canada
DIT	Double Index Tag
ER	Exploitation Rate
ERA	Exploitation Rate Analysis
ERI	Exploitation Rate Index
ESA	US Endangered Species Act
EV	Environmental Variable scalar
FI	Fishery Index
FNC	First Nations Caucus
FP	Fishery Policy Scalar

FR	Fraser River
FSC	Food, Social, and Ceremonial
GMR	Genetic Mark–Recapture
GW	Gitwinksihlkw
HR	Harvest Rate
HRI	Harvest Rate Index
iid	Independent Identically Distributed
IM	Incidental Mortality
ISBM	Individual Stock Based Management
JDF	Juan De Fuca
LAT	Low Abundance Threshold
LC	Landed Catch
LGS	Lower Strait of Georgia
LIM	Legal Incidental Mortality
MOC	Mid-Oregon Coast
MR	Mark–Recapture
MRE	Mature-Run Equivalent
MSE	Mean Squared Error
MSF	Mark-Selective Fishery
MSY	Maximum Sustainable Yield for a stock,
NA	Not Available
NBC	Northern B.C. Dixon Entrance to Kitimat including Haida Gwaii
NBC T	North British Columbia Troll
NC	North Coastal
NM	Nautical Mile
NMFS	National Marine Fisheries Service
NOC	North Oregon Coast
NWIFC	Northwest Indian Fisheries Commission
ODFW	Oregon Department of Fish & Wildlife
ORC	Oregon Coast
PNV	Proportion non-vulnerable
PS	Puget Sound
PSC	Pacific Salmon Commission
PST	Pacific Salmon Treaty
PT	Pre Terminal
PV	Proportion vulnerable
QCI	Haida Gwaii (Queen Charlotte Islands)
QIN	Quinault Indian Nation
RE	Relative Error

ROM	Ratio of Means
SA	Simple Average
SAFT	Stock, Age, Fishery and Time Period
SEAK	Southeast Alaska Cape Suckling to Dixon Entrance
SEAK T	SE Alaska Troll
SIM	Sublegal Incidental Mortality
SMSY	Escapement producing MSY
SPFI	Stratified Proportional Fishery Index
SPS	South Puget Sound
SSP	Sentinel Stocks Program
SUS	Southern US
TAC	Total Allowable Catch
TAC	*U.S. v Oregon* Technical Advisory Committee
TBR	Transboundary Rivers (Alsek, Taku, Stikine)
TLA	Three Letter Acronym
TM	Total Mortality
UAF	University of Alaska Fairbanks
UGS	Upper Strait of Georgia
UMSY	Exploitation Rate at MSY
UMT	Upper Management Threshold
URB	Columbia Upriver Brights
US	United States
USFWS	U.S. Fish & Wildlife Service
VB	Visual Basic
WA/OR	Ocean areas off Washington and Oregon North of Cape Falcon
WAC	Washington Coast
WCVI T	West Coast Vancouver Island Troll
WCVI	West Coast Vancouver Island excluding Area 20
WDFW	Washington Department of Fish and Wildlife