



SUSTAINABLE
SALMON
INITIATIVE

821 N Street, Suite 103, Anchorage, AK 99501 / PHONE (907) 279-6519 / FAX (907) 258-6688 / www.aykssi.org

**ARCTIC-YUKON-KUSKOKWIM SUSTAINABLE
SALMON INITIATIVE
INDEPENDENT EXPERT PANEL**

Review of:

**KUSKOKWIM RIVER CHINOOK SALMON RUN-
RECONSTRUCTION
&
STOCK-RECRUITMENT MODELS**

January 10, 2019

INDEPENDENT EXPERT PANEL
Review of
KUSKOKWIM RIVER CHINOOK SALMON RUN-RECONSTRUCTION &
STOCK-RECRUITMENT MODELS

Commissioned by:
Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative
821 N Street., Suite 103
Anchorage, AK 99501

AYK SSI INDEPENDENT EXPERT PANEL

Daniel Schindler (Chair), Professor, University of Washington, School of Aquatic and Fishery Sciences

Timothy Walsworth (Analyst), Post-Doctoral Researcher, University of Washington, School of Aquatic and Fishery Sciences

Milo Adkison, Professor, College of Fisheries and Ocean Sciences, University of Alaska Fairbanks

Randall Peterman, Professor Emeritus, School of Resource and Environmental Management, Simon Fraser University

André Punt, Professor, University of Washington, School of Aquatic and Fishery Sciences

This document should be cited as:

Schindler, D.E., T.E. Walsworth, M.D. Adkison, R.M. Peterman, and A.E. Punt. 2019. Kuskokwim River Chinook salmon run-reconstruction and stock recruitment models: a review by an independent expert panel. Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative. Anchorage, AK. 164 pp.

Executive Summary

A four member independent expert panel, assisted by a biometric analyst, was commissioned by the [Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative](#) to examine the performance of the Alaska Department of Fish & Game (ADFG) run-reconstruction model used for estimating the abundance of Chinook salmon in the Kuskokwim River and to provide recommendations for model improvements and future analysis. The model we evaluated (referred to as the 'original model' throughout this report) was used by ADFG to estimate Chinook salmon abundance in the Kuskokwim River until April 2018. We used two approaches to assess the performance of this model. First, we examined model performance using monitoring data collected between 1976 and 2017 to determine the influence of different input data (i.e., mark-recapture estimates) and various structures of abundance-estimation models on estimates of total run size. Second, we used simulated population data, intended to approximate the dynamics of Kuskokwim River Chinook salmon, to assess model performance when true values of model parameters and structures were known, as a way to examine model sensitivity (accuracy and bias) to violation of different model assumptions.

The original run-reconstruction model was highly sensitive to starting values used for abundance-related parameters; the estimation method found multiple local minima, demonstrating that it was not converging properly on a unique solution. The differences between abundance estimates across model fits were often large enough to be of management concern. The model convergence problem was reduced when over-dispersion parameters were pooled by type of escapement index (i.e., for data from weirs as one group, and from aerial surveys as another), and when the harvest component of the likelihood was removed. The influence of the harvest component on model stability was unexpected and should be explored further. We recommend examining model fits and negative log-likelihoods across a wide range of starting parameter values to ensure that the true global minimum negative log-likelihood is found. The complexity of the original model is not supportable with the data used to fit it because there are too many parameters relative to sources of information to estimate them. Further, the model assumption that errors in weir counts and aerial surveys are distributed according to a negative binomial distribution should be revised to assume that they are log-normally distributed.

Mark-recapture estimates of total in-river abundance are critical for interpreting abundance indices collected at the tributary scale across the river. Inclusion of the most recent (years 2014-2017) mark-recapture estimates reduced estimates of the run size of Kuskokwim Chinook salmon produced by the original model in the most recent half of the time series. In particular, inclusion of 2014-2017 mark-recapture studies produced lower abundance estimates for the recent time frame during which Chinook salmon abundance has been particularly low. Periodic mark-recapture estimates are critical for minimizing potential bias in run size estimates in the future, particularly when population dynamics are not synchronous across the river basin, and when the stock responds to changing environmental regimes, as appears to have occurred in the last decade.

The frequency at which mark-recapture studies should be conducted should be evaluated quantitatively using a formal value-of-information (VOI) approach based on simulations similar to those used by this panel. Such VOI analyses will determine whether and how often the large expense of mark-recapture programs is justified, in light of benefits such as the value of the additional catches or reducing the

probability of overharvest. This same VOI approach should also be used to determine the value of weir and aerial survey programs, across a range of levels of spatial coverage, because it is probable that individual enumeration projects will become increasingly important as sampling coverage of the watershed declines, particularly in situations where the stock components do not show synchronous dynamics across the river basin, as is apparent for Chinook salmon in the Kuskokwim River.

Performance of the original run-reconstruction model was only modestly sensitive to its assumptions that there are no regime shifts in oceanographic conditions affecting Chinook salmon productivity, and that there are synchronous population dynamics among sub-stock components. Analyses of simulated data demonstrate that errors in estimates (imprecision and bias) increased when sub-stock dynamics were not synchronous across the river, particularly in the presence of environmental regime shifts that cause large changes in the productivity of the stocks. However, pooling over-dispersion parameters in the original model (or variances in log-normal models) greatly reduced the impacts of regime shifts in asynchronous sub-populations on the bias in run estimates (i.e., when the model was misspecified). Pooling had much less of an effect when the sub-populations were synchronous.

Simulations showed that regime shifts producing low frequency changes in population productivity have the potential to make population dynamics appear as though they are determined by Ricker stock-recruitment dynamics (i.e., there is over-compensation in recruitment at high spawner densities) even though the simulated data are generated from Beverton-Holt models, which do not exhibit overcompensation. The tendency of regime shifts to produce apparent overcompensation was strongest when productivities were markedly different from one regime to another (> 4-fold difference in productivity), and when the frequency of regime shifts was roughly aligned with the longevity of the fish (~4-8 years). Whether the dynamics of Kuskokwim Chinook salmon are produced by overcompensation in the stock-recruitment relationship and/or from changing environmental regimes is not clear at present, but should be investigated further. Given these uncertainties, management strategies should be adopted that are robust to the possibility that either scenario is the true state of nature.

Simulations showed that assuming that population dynamics were determined by Ricker (over-compensatory) dynamics when, in fact, they were generated with Beverton-Holt (compensatory) dynamics in the presence of environmental regime shifts, led to more biologically conservative management benchmarks than if the model incorporated a Beverton-Holt model. Management objectives that seek to achieve maximum sustainable yield (MSY) that are supported by analyses assuming Ricker dynamics would have escapement goals that were > 2 times higher than the escapement goals based on Beverton-Holt models. Thus, because the underlying cause of apparent overcompensation is not currently known, assuming Ricker dynamics leads to more biologically precautionary management than assuming Beverton-Holt dynamics. Such biologically precautionary management that derives from implicitly assuming Ricker dynamics does come with the cost of reduced harvest opportunity, but should also provide protection to weak substocks within the Kuskokwim stock complex, which is characterized by asynchronous dynamics among its component populations. Future analyses should more fully explore the trade-offs among a variety of performance metrics in an attempt to understand the consequences of regime shifts and weak mechanistic understanding of the causes and strengths of the anomalously strong density-dependence in Kuskokwim Chinook salmon when compared with other Alaska Chinook salmon stocks.

Estimates of abundance from a 'revised model' adopted by ADFG in May 2018 (with log-normal errors on weir and aerial survey data, pooled variances within survey types, a reformulated harvest model, and integration of the most recent mark-recapture estimates) produced historical abundance estimates that tended to be lower than those produced by the original model, particularly in the recent decade. Using the revised model estimates as input to a stock-recruitment analysis showed that escapement abundance required to achieve MSY (or within 10% of it) would change little from targets based on the original model. However, abundance estimates from the revised run-reconstruction model suggest that the population is about 20% less productive at S_{msy} (the escapement that would produce long-term maximum sustainable yield) than inferred from output from the original model.

There is also another important trend to consider. Current observations of the increasing rarity of the largest and oldest fish in Kuskokwim Chinook salmon raise concern that the 'escapement quality' is declining through time, which could contribute to a decline in average per-capita productivity of the stock. Long-term changes in various aspects of escapement quality (e.g., size-at-age of the spawners, age composition, sex ratio) should be quantified given existing data. Future work should focus on quantifying the potential effects of these (and future) changes on the productivity of the stocks, and correspondingly, on relevant management reference points using simulation and analysis. Simulation should be used to explore the degree to which management reference points should be adjusted to account for changes in stock productivity in response to environmental variation and erosion of escapement quality, given the uncertainties in the data and the limited ability of the current monitoring program to detect future changes in escapement quality and run size.

Use of the stock-recruitment relationship to establish management reference points currently assumes a homogeneous stock of Chinook salmon distributed across the Kuskokwim river when, in fact, there is considerable evidence that Kuskokwim Chinook salmon are a collection of smaller stocks with some degree of independence in their dynamics. Management strategies that assume a homogeneous stock run the risk of overexploiting smaller and less productive stocks, thereby eroding the resilience of the overall stock complex to future changes in the environment. Thus, it is important to examine the consequences of various harvest regimes for sustainable production over the long-term, as well as maintaining the biocomplexity within the system. Maintaining biocomplexity in the system may require lower harvest rates in the short-term, and this trade-off should be explicitly examined using simulation and analysis of the existing data.

Any further refinements of the run-reconstruction and stock-recruitment models should be thoroughly tested through a simulation approach such as done here, to ensure that model behavior and robustness to inevitable violations of assumptions are thoroughly understood. Such testing of the model will also promote a more transparent model development process.

List of Figures

- Figure 1. Example simulations of sub-population dynamics under different production scenarios generated by the operating model.
- Figure 2. Run-reconstructions fit to actual data under different model structures (individual or pooled over-dispersion parameters) and with different mark-recapture data available (no new mark recapture estimates, all new mark-recapture estimates, without 2017 mark-recapture estimate).
- Figure 3. Changes in annual run-reconstruction estimates when switching from the base model (individual over-dispersion parameters and no new mark-recapture estimates available) to different model structures and data scenarios.
- Figure 4. Run estimates across a range of starting values from the original ADFG R model implemented in R without new mark-recapture estimates.
- Figure 5. Run estimates across a range of starting values from the original ADFG R model implemented in ADMB without new mark-recapture estimates.
- Figure 6. Run estimates across a range of starting values from the original ADFG R model implemented in R with new mark-recapture estimates.
- Figure 7. Run estimates across a range of starting values from the original ADFG R model implemented in ADMB with new mark-recapture estimates.
- Figure 8. Negative log-likelihood values across a range of parameter starting values under different model structures.
- Figure 9. Run estimates across a range of starting values from the ADFG R model implemented in R without new mark-recapture estimates and with over-dispersion parameters pooled by escapement index type.
- Figure 10. Run estimates across a range of starting values from the ADFG R model implemented in ADMB without new mark-recapture estimates and with over-dispersion parameters pooled by escapement index type.
- Figure 11. Run estimates across a range of starting values from the ADFG R model implemented in R with new mark-recapture estimates and with over-dispersion parameters pooled by escapement index type.
- Figure 12. Run estimates across a range of starting values from the ADFG R model implemented in ADMB with new mark-recapture estimates and with over-dispersion parameters pooled by escapement index type.
- Figure 13. Run estimates across a range of starting values from the ADFG R model implemented in ADMB without new mark-recapture estimates, and without the harvest component of the likelihood.
- Figure 14. Run estimates across a range of starting values from the ADFG R model implemented in ADMB with new mark-recapture estimates, and without the harvest component of the likelihood.

- Figure 15. Run estimates across a range of starting values from the ADFG R model implemented in ADMB without new mark-recapture estimates, and without the harvest component of the likelihood. Over-dispersion parameters are pooled by escapement index type.
- Figure 16. Run estimates across a range of starting values from the ADFG R model implemented in ADMB with new mark-recapture estimates, and without the harvest component of the likelihood. Over-dispersion parameters are pooled by escapement index type.
- Figure 17. R model in ADMB fit to simulated data without the harvest component of the likelihood. This scenario had new mark-recapture estimates incorporated, perfect synchrony among sub-stocks, and no observation error.
- Figure 18. R model in ADMB fit to simulated data without the harvest component of the likelihood. This scenario had new mark-recapture estimates incorporated, perfect synchrony among sub-stocks, and with observation error.
- Figure 19. R model in ADMB fit to simulated data without the harvest component of the likelihood. This scenario had new mark-recapture estimates incorporated, asynchronous sub-population dynamics, and no observation error.
- Figure 20. R model in ADMB fit to simulated data without the harvest component of the likelihood. This scenario had new mark-recapture estimates incorporated, asynchronous sub-population dynamics, and with observation error.
- Figure 21. Correlations between observed escapement indices and model predicted escapement indices across different estimation model structures.
- Figure 22. Escapement index over-dispersion parameters fit across a range of parameter starting values and model structures.
- Figure 23. Effects of fixing catch and effort data pooling across final three weeks of season on run-reconstruction estimates with individual over-dispersion parameters.
- Figure 24. Effects of fixing catch and effort data pooling across final three weeks of season on run-reconstruction estimates with pooled over-dispersion parameters.
- Figure 25. Change in run-reconstruction estimates from no longer pooling catch and effort data across the final three weeks of the season.
- Figure 26. Boxplots of normalized RMSE for Kuskokwim run-reconstruction model fits to simulated data of Kuskokwim River Chinook salmon under different biological scenarios and model structures. Compares the effects of asynchrony, new mark-recapture estimates, over-dispersion parameterization and harvest likelihood on normalized root mean squared error.
- Figure 27. Boxplots of normalized RMSE for Kuskokwim Run-reconstruction model fits to simulated data under different biological scenarios and model structures. Compares the effects of asynchrony, new mark-recapture estimates, over-dispersion parameterization and regime shifts on normalized root mean squared error.
- Figure 28. Boxplots of normalized RMSE for Kuskokwim Run-reconstruction model fits to simulated data under different biological scenarios and model structures. Compares the effects of

asynchrony, new mark-recapture estimates, over-dispersion parameterization and non-linear aerial bias on normalized root mean squared error.

- Figure 29. Absolute values of relative error through time in run-reconstruction estimates for simulated time-series where the harvest component of the likelihood was included.
- Figure 30. Absolute values of relative error through time in run-reconstruction estimates for simulated time-series where the harvest component of the likelihood was excluded.
- Figure 31. Absolute values of relative error through time in run-reconstruction estimates for simulated time-series with unique over-dispersion parameters for each escapement index.
- Figure 32. Absolute values of relative error through time in run-reconstruction estimates for simulated time-series with over-dispersion parameters pooled by escapement index type.
- Figure 33. Absolute values of relative error through time in run-reconstruction estimates for simulated time-series without non-linear bias in aerial survey indices.
- Figure 34. Absolute values of relative error through time in run-reconstruction estimates for simulated time-series with non-linear bias in aerial survey indices.
- Figure 35. Density plots of the absolute value of the average time-series bias across all iterations for different scenarios for both the operating model and estimation model, where the harvest component of the likelihood is included in the estimation model.
- Figure 36. Density plots of the absolute value of the average time-series bias across all iterations for different scenarios for both the operating model and estimation model, where the harvest component of the likelihood is excluded from the estimation model.
- Figure 37. Density plots of the absolute value of the average time-series bias across all iterations for different scenarios for both the operating model and estimation model, where the over-dispersion parameters in the estimation model were unique for each escapement index.
- Figure 38. Density plots of the absolute value of the average time-series bias across all iterations for different scenarios for both the operating model and estimation model, where the over-dispersion parameters in the estimation model were pooled by escapement index type.
- Figure 39. Density plots of the absolute value of the average time-series bias across all iterations for different scenarios for both the operating model and estimation model, where there is no non-linear bias in aerial survey indices.
- Figure 40. Density plots of the absolute value of the average time-series bias across all iterations for different scenarios for both the operating model and estimation model, where there is non-linear bias in aerial survey indices.
- Figure 41. The proportion of simulations in which the true run-size value was within the 95% confidence interval of the model estimated run size across time under different scenarios.
- Figure 42. Median estimates of the Shepherd's β parameter (as a measure of over-compensation), when the Shepherd spawner-recruit relationship is fit to estimates from the run-reconstruction model under different productivity regime and sub-population synchrony scenarios.

- Figure 43. Median estimates of the harvest rate that would produce MSY derived from a Ricker production function fit to run-reconstruction estimates relative to that which produced the maximum yield in simulated populations.
- Figure 44. Median estimates of the spawning escapement that would produce MSY derived from a Ricker production function fit to run-reconstruction estimates relative to that which produced the maximum yield in simulations.
- Figure 45. Comparison of Ricker stock-recruitment relationships for Chinook salmon in the Kuskokwim River fit to output from three run-reconstruction model scenarios. Top panel show output from each of the run-reconstruction model scenarios and the best-fit Ricker function fit. In the bottom panel shows the expected surplus production as a function of spawner abundance for each model scenario.

List of Appendices

Appendix 1. Original scope of work from the Steering Committee of the Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative (AYK SSI), commissioning this report.

Appendix 2. Review of ADFG methodology used in mark-recapture studies to estimate abundance of Chinook salmon in the Kuskokwim River; by Dr. John Skalski, University of Washington.

Appendix 3. Memo from AYK SSI Expert Panel to ADFG following preliminary review of revised run-reconstruction model following collaborative workshop between Expert Panel and ADFG to discuss current model limitations. Dated May 10, 2018.

Appendix 4. Computer code (R) for operating model used to generate simulated data for assessing performance of run-reconstruction models.

Appendix 5. ADMB code for original run-reconstruction model ('R model').

Appendix 6. Biographies of members of the AYK SSI Expert Panel.

Introduction

The Kuskokwim River is the second largest river in Alaska, draining a remote and vast watershed that serves as spawning and nursery habitat for five species of anadromous Pacific salmon, which support subsistence, commercial and sport fisheries. Fisheries for salmon are focused primarily on chum and Chinook salmon as they return as maturing adults to spawn throughout the watershed. Both species have shown large swings in abundance over the last few decades, which have challenged both the stakeholders and the managers of these resources. Management is hampered by the difficulties in performing accurate stock assessments needed to establish harvest regulations, which are reliant on data that are costly to collect from this remote and complex river where funding for sampling and assessment is distinctly limited.

The Chinook salmon stock in the Kuskokwim River is one of the largest wild populations in the world, averaging around 250,000 fish between 1976 and 2017. As with other populations in western Alaska (ADFG 2013, Schindler et al. 2013, Ohlberger et al. 2016), Kuskokwim Chinook salmon have shown precipitous declines in abundance and productivity over the last decade, with little scientific explanation. Though sparsely populated, people living throughout the watershed have had to cope with restrictions on subsistence fishing opportunities, and directed commercial fisheries for Chinook salmon were closed in 1987, though incidental harvest still occurs during commercial fisheries directed at chum and sockeye salmon. Reduced abundance of Chinook salmon has focused attention on the management process, with questions about the reliability of abundance estimates, how uncertainty in these estimates is incorporated into management decision-making, and whether management is sufficiently precautionary to protect the long-term sustainability of the stock at both the drainage-wide and tributary level.

Chinook salmon in the Kuskokwim River are managed based on an escapement goal policy (Hamazaki et al. 2012; Conitz et al. 2015) where harvest is restricted to achieve an abundance of adult fish on the spawning grounds that will sustain the population over the long term. Sustainable escapement goals are defined as a level of escapement, indicated by an index or an escapement estimate, that is known to provide for sustained yield over a 5- to 10-year period. These goals are used in situations where data limitation precludes the estimation of escapement that would produce maximum sustainable yield (MSY). The escapement level most likely to produce maximum sustainable yield (S_{msy}) is estimated from a stock-recruitment model that relates the number of spawners to the number of recruits produced in the next generation. Escapement that produces maximum sustainable yield is typically found at intermediate to low spawning densities where there is little competition for high-quality spawning sites among adults, and for food resources among juveniles. Fisheries are prosecuted such that the number of fish in an annual spawning run that exceeds the escapement goal are allocated for harvest.

Accurate annual estimates from the Kuskokwim Chinook salmon run-reconstruction model are also important because they help guide both in-river management and management of Chinook bycatch in marine fisheries in the Bering Sea. In April 2015, the North Pacific Fishery Management Council (Council) took action to lower Chinook salmon bycatch caps in the Bering Sea pollock fishery when the estimated Chinook salmon abundance falls near or below a low abundance threshold. By October of each year, ADFG must provide to the Council their '3-system index' consisting of the combined in-river adult Chinook salmon run sizes from the Unalakleet, Upper Yukon, and Kuskokwim rivers. This total estimate is then compared to the Council's identified threshold of 250,000 fish to determine if additional bycatch reduction actions by the pollock fleet are required. While estimates of the abundance of Chinook salmon

on the Unalakleet and Upper Yukon rivers are derived largely from direct counts, the Kuskokwim River annual estimate relies on the more complex run-reconstruction model to estimate abundance at the basin-wide scale.

Estimates of Chinook salmon abundance in the Kuskokwim River produced by the run-reconstruction model also play a key role in determining before the fishing season whether ADFG or USFWS (under delegation of authority from the Federal Subsistence Board) will serve as the lead in-season manager of the Chinook salmon run within federal waters on the Kuskokwim River. The vast majority of the subsistence harvest occurs in the lower portion of the river within the Yukon Delta National Wildlife Refuge. The Kuskokwim River Chinook salmon preseason forecast is currently based on the run-reconstruction model estimate of abundance in the previous year. If this estimate indicates that there will be insufficient Chinook salmon to meet both escapement and subsistence needs within the watershed and, therefore, that harvest must be restricted, the federal subsistence program may take action to limit harvest to federally qualified users under section 804 of the Alaska National Interest Lands Conservation Act (ANILCA). Due to depressed runs, action was taken by the Federal Subsistence Board in 2014-18 to federally manage the Chinook fishery within Yukon Delta NWR.

Estimating the escapement most likely to produce maximum sustainable yield (S_{msy}) is a highly technical process that relies on reliable estimates of Chinook salmon abundance through time, the distribution of returning fish between harvest and escapement and, in some cases, environmental effects on the recruitment process. In principle, this involves the simple process of counting the number of fish caught in fisheries and the number that survive to reach the spawning grounds. However, in practice, this is an extremely difficult endeavor, particularly in systems as large and remote as the Kuskokwim River where a full census of the population is impossible and indicators of abundance are sampled only patchily across space and sporadically through time. Even estimating the number of fish captured in fisheries is highly uncertain because the bulk of the catch is for subsistence purposes, and there is no centralized location where fish are processed and easily enumerated.

Given these difficulties of sampling Kuskokwim Chinook, statistical models are used to estimate the total abundance of fish that return to the river each year. These 'run-reconstruction' models are essentially mathematical ledgers intended to be a parsimonious representation of the harvest and enumeration process as fish return from the ocean and migrate throughout the river to spawn. Run-reconstruction models integrate observations that are indicators of abundance; i.e., the number of fish caught in fisheries, the number passing through weirs located on a subset of tributaries, and the number of spawners on a subset of spawning reaches observed in aerial surveys. All of these indicators are subject to substantial uncertainties, including both imprecision and bias, so uncertainties must be accounted for explicitly when estimating total run size. Mark-recapture experiments are also used to estimate total abundance because they provide an objective way to scale from tributary indicators of abundance to estimates of absolute total abundance. However, mark-recapture studies are expensive and are only performed periodically. Nonetheless, as we show below, they are critically important to the run-reconstructions because they provide the only reliable information about the absolute river-wide abundance of fish, thereby anchoring the indicators of relative abundance obtained via other observations.

The run-reconstruction model used by ADFG to estimate total abundance of Chinook salmon in the Kuskokwim River integrates observations from weirs, aerial surveys, fishery catches, and occasional mark-

recapture experiments. The primary challenge lies in relating observations among locations, methods, and years because the data are sparse. As described below, information from harvest, weirs, and aerial surveys are weighted equally in the run-reconstruction model used by ADFG to estimate total annual abundance of adult Chinook. The model also requires a number of simplifying assumptions to make the estimation problem tractable. However, the performance of the run-reconstruction model, its robustness to the underlying assumptions, and sensitivity to data quality have not yet been evaluated systematically.

An Independent Review Panel (henceforth "the Panel") was established by the Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative (AYK SSI) in October 2016 to investigate some of the challenges described above. The Panel's main objective was to systematically test the ADFG run-reconstruction and spawner-recruit models for Kuskokwim River Chinook salmon, assess their limitations in light of the types of data and their uncertainties available for informing the model, test the influence of the key assumptions, and recommend ways to improve the models in the future. Throughout this report we use the phrase '**original model**' to reference the ADFG run-reconstruction model that was used by the department until May, 2018. We use the term '**revised model**' to reference the run-reconstruction model that was revised to accommodate a number of suggestions from the Panel at a collaborative workshop held in March 2018 between the Panel and the ADFG model development team. As described below, the revised model includes a different error structure and different harvest model from the original model, that appeared to remedy most of the problems identified by the Panel in its review of the original model (see Appendix 3).

The Panel's research questions were guided by some chief concerns about the run-reconstruction model that were either reported by ADFG or were raised by stakeholders and previous explorations of the model. Those concerns were the following.

- The first major concern was that the model estimates of total run abundance appeared to be sensitive to the starting parameter values, an indication that either there was a structural flaw in the model, its parameterization was inappropriate, the data were not informative concerning abundance, or the method used for parameter estimation was not performing as expected.
- Second, the original model assumed that the proportional returns of Chinook salmon to the different tributaries throughout the Kuskokwim River are perfectly synchronous through time. This assumption is tenuous given the expanding literature documenting the diversity in population dynamics of salmon at regional and within-watershed scales. Although there is evidence of positive correlation in productivities among salmon populations, including Chinook salmon, there also is residual variation that is not shared among populations at one or both of those scales (e.g., Peterman et al. 2003; Pyper et al. 2005; Dorner et al. 2008, 2018; Rogers and Schindler 2008; Schindler et al. 2010; Sharma et al. 2013; Kilduff et al. 2014). However, the sensitivity of the run-reconstruction model to the assumption of perfect synchrony has not been assessed.
- Third, the original model assumed that production dynamics of Chinook salmon in the Kuskokwim are stationary through time, i.e., that changes in abundance are driven by a relationship between the spawning stock and recruitment whose parameters do not vary through time. Again, the assumption of stationarity is tenuous given the substantial empirical evidence that salmon populations, including those of Chinook salmon, routinely demonstrate non-stationarity (i.e., change in mean and/or variance) in their dynamics (Adkison et al. 1996; Peterman et al. 1998,

2003; Pyper et al. 2005; Dorner et al. 2008, 2018; Ohlberger et al. 2016; Peterman and Dorner 2012; Kilduff et al. 2014; Malick and Cox 2016). The influence of such non-stationarity on performance of the Kuskokwim run-reconstruction model has not been assessed systematically.

- Fourth, the original model relied heavily on abundance metrics derived from aerial surveys of fish on the spawning grounds. While these data are markedly cheaper to collect than data from weirs and mark-recapture studies, they are well-known to be highly uncertain, are usually biased, and their influence on the run-reconstruction model were unexplored.
- Fifth, the influence of mark-recapture data on the run-reconstruction model output has not been evaluated. Mark-recapture estimates of abundance are the only system-wide estimate of absolute abundance and therefore inform most other parameters in the model by providing a scalar to relate the relative abundance indices to absolute abundance. The value of recent mark-recapture estimates of abundance (2014-2017) for affecting model performance has not been assessed, particularly in light of the concerns described above.
- Finally, the stock-recruitment relationship implied by the model's abundance estimates demonstrates very strong density-dependence (strong overcompensation at high spawner densities, in fact) and suggests that escapement goals should be reduced if maximizing yield is the goal of management, an action that has poorly understood consequences if the estimated relationship were incorrect.

In response to this list of concerns, and from some general interests of the Panel, analyses were structured around the questions listed below. We used two primary approaches in our model assessment. First, we evaluated how sensitive the original run-reconstruction model was in terms of its estimates of total abundance of Chinook salmon when analyzing the existing data for the Kuskokwim River. Second, we used a simulation approach to assess model performance where an 'operating model' was constructed to simulate reasonable approximations of the population dynamics of Chinook salmon in the Kuskokwim River, but where we could specify the critical parameters and assumptions. The data resulting from these simulations were then 'sampled' and interpreted via the original run-reconstruction model to assess its ability to accurately infer the true underlying dynamics of the simulation. This approach provides an objective way to assess the biases and limitations of the original run-reconstruction model across a range of conditions that reflect different possibilities about the true nature of the dynamics of Chinook salmon in the Kuskokwim River.

The Panel's Main Questions

1. What modifications to the original run-reconstruction model (i.e., estimation model) could alleviate starting value sensitivity? Are there structural changes to the model and its associated estimation scheme that would make it more stable and insensitive to starting conditions?
2. What are the effects of the following assumptions on the run-reconstruction model's performance?
 - a. Synchrony in the productivity parameter that affects residuals in $\log_e(R/S)$ from the underlying spawner-recruit model among stocks in different spawning tributaries
 - b. Temporal autocorrelation in residuals of $\log_e(R/S)$ and regime shifts in population productivity parameters (Ricker α or Beverton-Holt α/β)
 - c. The non-linearity of biases in aerial survey indices of escapement
3. How does the availability of different data types affect run-reconstruction model performance?
 - a. What is the influence of mark-recapture data on model performance?
 - b. Does the influence of mark-recapture data interact with underlying recruitment dynamics (i.e. synchrony, temporal autocorrelation, and existence of regime shifts) to affect model performance?
4. How does the parameterization of the run-reconstruction model affect its performance?
 - a. Does estimating individual vs. pooled escapement over-dispersion parameters lead to better estimation performance and stability?
 - b. What are the effects on model behavior and performance of the original assumption of a negative-binomial error distribution on weir and aerial survey abundance indices compared to an alternative assumption of log-normally distributed errors?
 - c. Is the component of the model that estimates harvest properly formulated and integrated into the likelihood calculations?
5. What are the potential management implications of errors produced by the concerns raised above in terms of:
 - a. The model estimate of escapement that produces MSY (S_{msy}) from analyses of spawning stock size and subsequent recruitment?
 - b. The lost potential harvest rate in the system from mis-specifying the true S_{msy} ?
6. If regime shifts in recruitment productivity produce abundance dynamics that appear to be driven by strong density-dependent regulation (i.e., overcompensation),
 - a. What are the implications of such dynamics for estimating escapement that produces MSY (S_{msy})?
 - b. What are the implications for estimates of the sustainable harvest rates from mis-specifying the true S_{msy} ?

Approach

We used two approaches to assess the performance of the ADFG run-reconstruction model for Chinook salmon on the Kuskokwim River: (1) fitting the run-reconstruction model to the observed data sets supplied by ADFG, but with various modifications to that model's structure, and (2) fitting ADFG's original run-reconstruction model, as well as various modifications to it, to simulated data sets where the true parameter values and run sizes are specified and are intended to simulate alternative plausible states of

nature for the Kuskokwim River. A limitation of examining model performance on observed data is that the true state of the system is never known and there is no way to assess whether the model is actually capturing the true underlying dynamics in the system. Simulations allow for testing the model under different scenarios while being able to compare model fits to true values (Hilborn and Walters 1992).

Structure of the Estimation Model

The original run-reconstruction model used by ADFG for stock assessment relied on estimates of total in-river abundance from mark-recapture experiments to scale the abundance of Chinook salmon in the Kuskokwim River across the period of data collection, which includes many years without system-wide assessments of escapement. Relationships between the escapement indices at various monitoring projects and total escapement, as well as between harvest, effort and Chinook salmon abundance in years with information on total in-river abundance are used to estimate the abundance in years without direct estimates of total abundance. The run-reconstruction model incorporates information across all of these data sources in multiple sub-models that are combined into a single likelihood to determine the most likely parameter estimates, including annual estimates of total run size.

Run timing

Run timing was assessed using information from the Bethel test fishery, provided by the ADFG. The proportion of the Chinook salmon run present in week j and year y is assumed to be reflected by the catch rate in the fishery, as:

$$p_{yj} = \frac{CPUE_{yj}}{\sum_i CPUE_{yj}}, \quad (1.1)$$

where CPUE is the catch-per-unit-effort of the test fishery (Bue et al. 2012).

Escapement Indices

Escapement is monitored at multiple locations and by multiple methods in the Kuskokwim River basin. ADFG and the US Fish and Wildlife Service (USFWS) have maintained six weirs for varying lengths of time, with the earliest being started in 1976 (Kogruklu River). Weir counts were sparse before 2001, but much more consistent across all systems and years since 2001. However, only nine years have data from all six weirs available. Weir escapement counts typically monitor less than 20% of the estimated total run to the Kuskokwim River (Bue et al. 2012).

The original ADFG run-reconstruction model assumed that each weir enumerates a constant proportion of the total escapement to the Kuskokwim River each year (i.e., the relative runs to each tributary are perfectly synchronous through time), and thus that the expected weir count \hat{I} at site w in year y is related to total Chinook salmon escapement E in year y by:

$$\hat{I}_{wy} = \frac{E_y}{\hat{k}_w}, \quad (1.2)$$

where \hat{k}_w is the escapement index scaling factor for weir w .

ADFG has also conducted aerial escapement surveys on 14 tributary systems since 1976. As with the weir data, the aerial escapement data are patchily available through time. Only one year (2004) has data available from all 14 aerial index streams.

The run-reconstruction model assumed that, as is the case for weirs, each aerial survey index represents a constant proportion of the total escapement each year, and thus that the expected aerial survey index \hat{I} in at site a in year y is related to total Kuskokwim River Chinook salmon escapement E in year y by:

$$\hat{I}_{ay} = \frac{E_y}{\hat{k}_a}, \quad (1.3)$$

where \hat{k}_a is the escapement index scaling factor for aerial index location a .

In the original model, the likelihood of weir and aerial survey indices were assumed to be negative-binomially distributed around the expected indices estimated from the total escapement (E) and site specific scaling factors (k) according to:

$$f(I_{wy}; \hat{I}_{wy}; \hat{m}_w) = \frac{\Gamma(\hat{m}_w + I_{wy})}{\Gamma(\hat{m}_w) I_{wy}!} \left(\frac{\hat{I}_{wy}}{\hat{m}_w + \hat{I}_{wy}} \right)^{I_{wy}} \left(\frac{\hat{m}_w}{\hat{m}_w + \hat{I}_{wy}} \right)^{\hat{m}_w}, \quad (1.4)$$

$$f(I_{ay}; \hat{I}_{ay}; \hat{m}_a) = \frac{\Gamma(\hat{m}_a + I_{ay})}{\Gamma(\hat{m}_a) I_{ay}!} \left(\frac{\hat{I}_{ay}}{\hat{m}_a + \hat{I}_{ay}} \right)^{I_{ay}} \left(\frac{\hat{m}_a}{\hat{m}_a + \hat{I}_{ay}} \right)^{\hat{m}_a}, \quad (1.5)$$

where \hat{I}_{wy} is the expected (i.e., model-based) value of weir escapement index I_{wy} , \hat{m}_w is the over-dispersion parameter for weir site w , \hat{I}_{ay} is the expected value of aerial escapement index I_{ay} , and \hat{m}_a is the over-dispersion parameter for aerial survey site a (Bue et al. 2012).

Harvest

The harvest component of the estimation model relates weekly (j) harvest and effort data in the fishery to total estimated abundance by week (\hat{N}_{yj}) by estimating a catchability coefficient (q) for each of three different fishing regulation periods. Total estimated abundance by week was calculated as:

$$\hat{N}_{yj} = \hat{N}_y p_{yj}, \quad (1.6)$$

where \hat{N}_y is the estimated total run size for year y and p_{yj} is the proportion of the run returning the river in week i of year y .

The log-likelihood of catchability estimates were calculated as:

$$L(q, \hat{N}_y | F_{yj}, H_{yj}) = \prod_y \prod_j \frac{1}{\sigma_y \sqrt{2\pi}} \exp\left(\frac{\log(F_{yj}) - \log(\hat{F}_{yj})}{2\sigma_y^2}\right)^2, \quad (1.7)$$

$$\hat{F}_{yj} = -\log\left(1 - \frac{H_{yj}}{\hat{N}_{yj}}\right) / q$$

where F_{yj} is the fishing effort in year y and week j , H_{yj} is the fisheries harvest in year y and week j , and n is the total number of observations of harvest for catchability coefficient q . Separate q parameters were estimated for each of the three different regulation periods. As in Liller and Hamazaki (2017), the concentrated likelihood function was used to eliminate the need for estimation of variance for commercial efforts.

Total In-River Abundance

Total in-river abundance estimates were available from two separate mark-recapture projects, and these were used to anchor the rest of the run-reconstruction model so that it could estimate total abundance in years when only harvest, aerial surveys or weir data are available. One mark-recapture project operated from 2008 through 2012, and another produced estimates from 2014 through 2017. The latter four years are what we refer to in this report as "new" mark-recapture data. Until May 2018, only the 2008-2012 estimates were used by ADFG within the Chinook salmon run-reconstruction model due to concerns about comparability of the estimates from the early and later periods. The Panel's starting-parameter sensitivity analyses were run prior to the 2017 mark-recapture estimates becoming available and, as such, they were run without those estimates.

The negative log-likelihood of total run abundance was calculated as:

$$NLL(\hat{N}_y | R_y, \sigma_{ry}) = \frac{1}{2} \frac{(R_y - \hat{N}_y)^2}{\sigma_{ry}^2}, \quad (1.8)$$

where R_y is the run size estimate from the mark-recapture study in year y , and σ_{ry} is the standard deviation of the run size estimate from the mark-recapture study in year y .

The different likelihood components are given equal weight when the combined likelihood across all parameters was calculated.

Alternative Structures of the Estimation Model

The Panel examined three alternative estimation-model structures in addition to the ADFG original model described above. One alternative model, hereafter referred to as the "pooled" model, assumed that escapement indices of the same type (weir / aerial) had shared over-dispersion parameters. Instead of six \hat{m}_w parameters (one for each weir location) and 14 \hat{m}_a parameters (one for each aerial survey location), the pooled model has only one \hat{m}_w and one \hat{m}_a parameter that was shared among escapement indices of the same type.

The second alternative estimation model, hereafter referred to as the “no-harvest” model, ignored the harvest component of the likelihood (eqn. 1.7). The weekly effort data were not included in the data set; thus, the combined likelihood only included the total run and escapement component. We investigated this alternative estimation model because during certain analyses, we noted a strong effect of omitting harvest data, as we discuss later.

The third alternative estimation model combined the pooled and no-harvest models. This model will be referred to as the “pooled/no-harvest” model.

Operating Model Structure

Simulation experiments allow us to examine the potential biases produced by an estimation model by comparing its output across various parameter sets of an operating model that is parameterized to simulate different plausible states of nature. To run simulation experiments, we developed an operating model to generate (1) population dynamics that resemble the general patterns of the Kuskokwim Chinook salmon population, (2) commercial fishery effort and harvest, as well as (3) generate pseudo-observations of the system that are intended to emulate weir counts, aerial survey indices, and mark-recapture estimates of total in-river abundance (see Appendix 4 for computer code).

Population structure and dynamics

The number of Chinook salmon returning to the Kuskokwim system during year y , N_y , is given by

$$N_y = \sum_p \sum_a m_{p,y-a,a} R_{p,y-a} \quad (2.1)$$

Where $m_{p,y-a,a}$ is the proportion of animals of age a from population p that return during year y , $R_{p,y-a}$ is the total recruitment from population p for brood year y :

$$R_{p,y} = P_p(E_{p,y}) e^{\sigma_{rp} \varepsilon_{p,y} - \sigma_{rp}^2 / 2} \quad (2.2)$$

where $P_p(E_{p,y})$ is the stock-recruitment relationship for population p , σ_{rp} is the standard deviation of recruitment anomalies for population p , $\varepsilon_{p,y}$ is the recruitment residual for population p in year y :

$$\varepsilon_{y,p} = \rho \varepsilon_{y-1,p} + \sqrt{1 - \rho^2} \eta_{y,p} \quad \eta_{y,p} \sim MVN(0, \sigma_r^2) \quad (2.3)$$

where σ_r^2 is the variance-covariance matrix for drawing multivariate normal process errors for each population, and ρ is the magnitude of autocorrelation in recruitment anomalies. The variance-covariance matrix of sub-population recruitment anomalies were generated as:

$$\mathbf{COR} = \begin{bmatrix} 1 & \dots & z \\ \dots & 1 & \dots \\ z & \dots & 1 \end{bmatrix},$$

$$V = \begin{bmatrix} \sigma_{r1}\sigma_{r1} & \cdots & \sigma_{r1}\sigma_{rn} \\ \cdots & \cdots & \cdots \\ \sigma_{r1}\sigma_{rn} & \cdots & \sigma_{rn}\sigma_{rn} \end{bmatrix}, \quad (2.4)$$

$$\sigma_r^2 = \mathbf{COR} \circ V,$$

where z is the minimum correlation among sub-population productivity residuals, and n is the total number of sub-populations, and values noted by ellipses in **COR** are linearly interpolated between 1 and z down rows and columns. Values on the diagonal of **COR** are all equal to 1.

All sub-populations' recruitment dynamics were generated from either Beverton-Holt or Ricker stock-recruitment relationships:

$$P(E_p) = \begin{cases} \alpha_p E_p e^{-\beta_p E_p} & \text{Ricker} \\ \frac{\alpha_p E_p}{\beta_p + E_p} & \text{Beverton - Holt} \end{cases} \quad (2.5)$$

To simulate the effect of regime shifts on ecosystem productivity, two levels of productivity were generated for each population by setting productivity at the origin (α in Ricker relationship, α/β in Beverton-Holt relationship) to 10 for periods of high productivity, and to 5 for periods of low productivity. Maximum abundance of recruits for each sub-population (α_p for the Beverton-Holt, $\alpha_p/(\beta_p e)$ for the Ricker) was randomly drawn from a normal distribution $N(\text{mean}=25\ 000, \text{standard deviation}=6\ 000)$ to simulate populations of different sizes.

Throughout all simulations, we assumed a constant age-structure (19% total age 4, 38% total age 5, 39% total age 6, and 4% total age 7) across brood years and populations, informed by age-structure estimates in the Kuskokwim River Chinook populations.

Harvest Dynamics

Harvest was assumed to occur exactly as specified in the estimation model, where total harvest was calculated as:

$$H_{yjp_a} = N_{yjp_a} S_a (1 - e^{-q F_{yj}}), \quad (2.6)$$

$$N_{yjp_a} = N_{yp_a} p_{yjp}, \quad (2.7)$$

where H_{yjp_a} is the true harvest in year y week j of age a fish from sub-population p , N_{yjp_a} is the number of age a fish from sub-population p present in week j of year y , S_a is fishery selectivity for age a fish, q is the specified catchability coefficient, F_{yj} is the fishery effort in week j of year y , N_{yp_a} is the number of age a fish from sub-population p returning to the Kuskokwim during year y , and p_{yjp} is the proportion of all fish in population p in year y returning in week j . The operating model assumed a single fishery.

We varied annual fishing effort such that mean annual harvest rate was 0.49, and individual year harvest rates varied between ~ 0.05 and ~ 0.8 , with the distribution of values being negatively skewed (median

harvest rate = 0.53, sd = 0.21) due to the saturating relationship between effort and harvest rate (eqn. 2.6).

Observations

In the simulations, 6 of the 40 sub-populations were monitored for escapement using weirs, while 14 other sub-populations were monitored using aerial surveys. We assumed that all of the salmon entering the systems with weirs were counted, while aerial surveys counted only a proportion of the total number in the system.

$$I_{wy} = N_{py}e^{\varepsilon_w}, \quad (2.8)$$

$$I_{ay} = N_{py}Ae^{\varepsilon_a}, \quad (2.9)$$

where A is the proportion of the total run that is counted by aerial surveys, ε_w is the observation error on weir counts, and ε_a is the observation error on aerial surveys. This generates log-normal errors in escapement indices, which is different from those assumed in the estimation model (negative-binomial). This allows for smaller amounts of observation errors to be assessed, because the negative binomial distribution requires variance to be greater than the mean. As an alternative scenario, we assumed that aerial surveys were increasingly biased low as escapement increased:

$$I_{ay} = e^{0.59}N_{py}^{0.83}e^{\varepsilon_a} \quad (2.10)$$

where the exponents in equation 2.10 are from Jones et al. (1998).

Mark-recapture estimates of the total Chinook salmon in-river abundance were generated with:

$$R_y \sim LN(N_y, \sigma_{mr}), \quad (2.11)$$

where σ_{mr} was similar to the coefficient of variation of actual mark-recapture estimates in the given year (0.15 for original mark-recapture series, 0.05 for recent mark-recapture series).

Harvest and fishing effort were assumed to be observed with small observation errors (log-normal error 0.00001). This was a conservative approach towards assessing the utility of the harvest component of the model; in reality, there is probably substantial uncertainty in characterizing the subsistence fishery that dominates catch in this system. To simulate the irregularity of data availability across years, observations of escapement and total run were removed in all years for which there were no data available in the actual data set.

Simulation Scenarios

To examine the impacts of different assumptions on the performance of the estimation model, we generated data sets using the operating model that were intended to simulate alternative, but plausible, states of nature (Table 1). The primary assumptions that we examined were (a) the degree of synchrony among sub-populations, (b) the presence or absence of regime shifts in productivity, and (c) the presence

or absence of non-linearity in aerial survey indices. Additionally, we examined the impacts of various levels of data availability by running scenarios where only the old mark-recapture estimates were available, and scenarios where both the old and new estimates were available. Finally, we examined the accuracy and bias of the different estimation model structures (i.e., base model, pooled model, no-harvest model, and pooled/no-harvest model).

We characterized the different assumptions used in our analyses as follows. In scenarios where the sub-populations showed varying degrees of independence in their recruitment dynamics (i.e., were not perfectly synchronous with one other), the average correlation among residuals in $\log_e(R/S)$ of sub-populations was 0.29, the maximum correlation between a pair of stocks was 1, and the minimum was -1. Regime shifts were implemented at random, with a probability of occurring in each year of $(1/\text{regime frequency})$. A regime shift caused all populations to switch from a high-productivity phase to a low-productivity phase, or vice versa. Scenarios without aerial survey bias used equation (2.8) to generate aerial survey indices of escapement, while those that assumed bias in aerial surveys used equation (2.9). Finally, we assessed the performance of the estimation model under scenarios that included differing number of years of mark-recapture estimates that are used to inform the other parameters of the run-reconstruction model. Example sub-stock dynamics across different assumptions are shown in Fig. 1.

Table 1. Parameters and data sets that were varied under alternative scenarios to test the effect of violating assumptions on the accuracy and bias of the original run-reconstruction model.

| Parameter/ Data Type | Original Model Assumption | Alternative Parameterization |
|---|------------------------------|---------------------------------|
| Population synchrony (average correlation) | 1 | 0.29 |
| Regime-shift frequency | 0 | ~1/20 yr |
| Aerial survey bias | constant | Increases with N |
| Mark-recapture estimates | 2003-2007 | 2003-2007, 2014-2017 |

Model Diagnostics

Performance of the estimation model was quantified using several metrics. To assess the accuracy of annual run estimates for the entire Kuskokwim system, we calculated the normalized root mean square error (NRMSE):

$$NRMSE = \frac{\sum_y \sqrt{(\widehat{N}_y - N_y)^2}}{\bar{N}} \quad (3.1)$$

of model-estimated annual run size \widehat{N}_y relative to specified values from the operating model N_y . We normalized the root mean square error by the average run-size \bar{N} across all years to make model performance comparable across different stochastic iterations of the model.

We calculated the absolute value of the relative error and bias in run estimates as:

$$RE_y = \left| \frac{\hat{x}_y - x_y}{x_y} \right| \quad (3.2)$$

$$Bias_x = \frac{\sum_y \frac{\hat{x}_y - x_y}{x_y}}{y} \quad (3.3)$$

where x is the specified value of the parameter or variable (e.g., run size) of interest from the operating model output, and \hat{x} is the model estimate of the parameter from the estimation model.

We ran multiple iterations (n=500) for a subset of the model scenarios to examine the coverage probability of model estimates. Here we examined how often the simulated run size fell within the confidence interval for each year's run estimate under scenarios with either synchronous or asynchronous sub-population dynamics, with or without regime shifts, and with or without the recent mark-recapture data being available. All scenarios run in the coverage analysis had observation error incorporated, as well as non-linear bias in the aerial survey indices.

In addition to examining the accuracy and bias of parameter estimates, it was important to consider the management implications of such errors. To examine whether the estimates derived from the run-reconstruction produced apparent over-compensatory dynamics while being produced by Beverton-Holt production functions, we fit the Shepherd production function to the spawner and recruitment estimates derived from the run-reconstruction model:

$$\ln(R_t) = \ln\left(\frac{\alpha S_t}{1 + \left(\frac{S}{K}\right)^\beta}\right) + \varphi \omega_{t-1} + \varepsilon \quad (3.4)$$

$$\omega_t = \ln(R_t) - \mu_t$$

$$\varepsilon \sim N(0, \sigma_r)$$

where R is the number of recruits produced by spawners S , and α , β and K are estimated parameters, ω_{t-1} is the recruitment residual in time $t-1$ and φ is a measure of temporal autocorrelation in recruitment residuals. The shape of the Shepherd spawner-recruit curve depends on the value of β , where $\beta > 1$ produces over-compensatory dynamics, $\beta = 1$ produces a Beverton-Holt curve, and $\beta < 1$ produces only weak compensation in production dynamics. We examined the values of β estimated across a range of regime-shift strengths and durations to examine the conditions under which populations produced by Beverton-Holt dynamics would produce apparent over-compensation.

To examine the management implications of finding over-compensatory dynamics when they are not the true underlying driver, we first calculated \hat{u}_{msy} (the harvest rate that would produce maximum sustainable yield *MSY*) and \hat{S}_{msy} (the spawning escapement that would produce *MSY*) from a Ricker spawner-recruit relationship fit to brood table data generated from the estimation model output aggregated across the entire river basin. We used a Ricker stock-recruit function because this is the relationship used by ADFG to estimate the spawner-recruit relationship for Chinook salmon in the Kuskokwim River, and also because a Beverton-Holt relationship provided a poor fit to the aggregate (i.e., when sub-population structure is ignored) spawner-recruit data.

We fit the Ricker spawner-recruit relationship with autocorrelated residuals as follows:

$$\begin{aligned} \ln(R_t) &= \ln(\alpha) + \ln(S_t) - \beta S_t + \varphi \omega_{t-1} + \varepsilon \\ \omega_t &= \ln(R_t) - \mu_t \\ \varepsilon &\sim N(0, \sigma_r) \end{aligned} \tag{3.5}$$

where ω_{t-1} is the recruitment residual in time $t-1$ and φ is a measure of temporal autocorrelation in recruitment residuals (this equation is used in Hamazaki et al. 2012). We then simulated the sub-populations forward fifty years across a range of harvest rates (from 0 to 0.99, held constant across all fifty years) to determine the harvest rate (and escapement goal) that would produce *MSY*. We examined the error in the estimates calculated from the Ricker function to determine if assuming a Ricker spawner-recruit relationship causes management reference points to over- or under-harvest fish, and how the presence and characteristics of regime shifts affect any management reference points.

Starting-Parameter Sensitivity

To examine the sensitivity of the estimation model to starting values of total adult abundance (run size), a concern highlighted by ADFG staff (Hamazaki and Liller 2015), we fit the model across 100 different starting values for run size and examined the solutions produced by the estimation model. We considered a range of starting values from 100 000 to 400 000 (using the same starting value for all years), which brackets the previously estimated run sizes estimated for Chinook salmon in the Kuskokwim River. We examined the sensitivity of the model estimates to starting values using all four of the Panel's alternative model structures (Table 1) to determine why any minimization issues may have been occurring and to examine potential model-structure fixes to this issue. The starting-value issue was examined in the estimation model implemented in two modeling platforms, the R Statistical Computing Environment (R Core Team 2016) and AD Model Builder (ADMB; Fournier et al. 2012), as well as using both the actual Kuskokwim River Chinook salmon data and simulated data sets generated by the process model. For simulated data sets, we considered a range of starting values from 0.8 to 1.2 times the mean mark-recapture estimate of the simulated data set.

Results

Effects of Data-Availability on Run Size Estimates

Mark-recapture estimates from 2014-2017 were not included in official ADFG run-reconstruction analyses prior to May 2017 due to concerns about comparability with older mark-recapture estimates (e.g., estimates were made for different sections of river in different years, and used different methods to estimate downstream escapement). Stakeholders have raised concerns regarding the exclusion of recent mark-recapture estimates from the recent run-reconstruction analysis. For the analyses described in this report, ADFG supplied standardized estimates of mark-recapture for both the recent projects (2014-2017) and older projects (2003-2007), which allowed us to explore the effects of different amounts of data and alternative structures of the run-reconstruction model on the estimates of run-size in recent years. Both sets of mark-recapture estimates were standardized to the in-river abundance of Chinook salmon upstream of Birch Tree Crossing (river kilometer 294), plus estimates of unmonitored escapement to downstream tributaries. The original downriver escapement expansions were estimated from modeled relationships between watershed size and productivity using monitored reference tributaries to scale unmonitored tributaries (Schaberg et al. 2012). These expansions were found to be biased high (ADFG, *unpublished data*), and the new lower river expansions are on average 62% of the expansions reported in Schaberg et al. (2012), resulting in smaller total run estimates using the new expansions.

Regardless of whether new mark-recapture data were included or excluded, run size estimates followed qualitatively similar trends throughout the time series (Fig. 2). However, exclusion of the newest set of mark-recapture estimates resulted in slightly higher run-size estimates in recent years (2011-2017), regardless of whether over-dispersion parameters were pooled or unique for different escapement indices (Fig. 2). While there were consistent differences in point estimates of run size, estimates from all models fell within the confidence intervals of all other model estimates (Fig. 2), indicating no statistically significant difference between model runs. Run-estimate differences between the different model structures and data sets were greater in the early years of the data set than in later years (Fig. 3).

Starting-value Sensitivity in Estimation Model Using Observed Data

The original estimation model was sensitive to starting values of run size, arriving at multiple solutions, regardless of whether new mark-recapture estimates were included (Figs. 4-8). This issue was particularly problematic when fitting the model in ADMB (Fig. 5, 7, 8), but it was still an issue when fitting in R (Figs. 4, 6, 8). The average difference in annual run estimates across model solutions was 18,320 when fit in R (Fig. 4) and 75,670 when fit in ADMB (Fig. 5) without new mark-recapture estimates, while the maximum difference (red lines in Figs. 4-7) was 49,550 in R and 220,300 in ADMB. When new mark-recapture estimates were available, the differences among solutions across starting values improved slightly, with mean values of 11,470 in R (Fig. 6) and 57,940 in ADMB (Fig. 7) and maximum values of 55,650 in R and 195,500 in ADMB. However, while the range of solutions found by ADMB was much wider than in R, the best solutions found in ADMB had lower negative log-likelihood values than those found in R, indicating better fits to the data (Fig. 8). These multiple solutions in annual run-size values are a concern because they can have different management implications in terms of future decisions and interpretation of the effectiveness of past regulations. The different solutions, many of them found repeatedly by the estimation procedure, had different minimum negative log-likelihood values, suggesting the presence of

multiple local minima across the likelihood surface, potentially due to over-parameterization of the estimation model (Fig. 8). Further, the different repeated model solutions were arrived at across the entire range of starting values (Fig. 8).

To examine whether the apparent over-parameterization resulted from the unique over-dispersion parameters for each escapement project, we examined the effect of starting values on model estimates when using the pooled model, and when including or excluding the newest mark-recapture estimates of total escapement. Pooling the over-dispersion parameters into a single parameter for aerial surveys and another for weirs reduced the variability among model fits for most implementations of the model (Figs 9-12). The model with pooled over-dispersion parameters in R without the new mark-recapture estimates had a lower mean (16,290) and maximum (47,260) difference among run estimates (Fig. 9) than when over-dispersion parameters were not pooled (Fig. 4). The model fit in R with the new mark-recapture estimates was slightly more variable with pooled over-dispersion parameters (Fig. 11) than without (Fig. 6), with a mean difference among solutions of 14,500 and maximum difference of 54,920 (Fig. 11). The model fit in ADMB with pooled over-dispersion parameters (Fig. 10) was on average less sensitive to starting values than with individual over-dispersion parameters on average when the new mark-recapture estimates were ignored (Fig. 5) (mean difference among solutions of 57,820 and maximum difference of 223,900; Fig. 10). Additionally, the model in ADMB with pooled over-dispersion parameters was much more stable when the new mark-recapture estimates were incorporated (mean difference among solutions of 13,780 and maximum difference of 58,550; Fig. 12) than when each escapement index had a unique over-dispersion parameter (Fig. 7). Once again, while the model solutions in ADMB were more variable than in R, the best solutions in ADMB had lower negative log-likelihood values than in R, indicating better fits to the data (Fig. 8). A key finding is that sensitivity to starting values decreased the most when both over-dispersion parameters were pooled and the new mark-recapture data were used in the estimation.

Because there still was evidence of multiple local minima when we assumed pooled over-dispersion parameters and used all the mark-recapture data (Fig. 8), we examined the sensitivity of the estimation model to inclusion of the harvest component in the calculation of the negative log-likelihood. We hypothesized that the catchability coefficients during the early years of the time series might be causing fitting problems, because there are no anchoring mark-recapture estimates and few escapement indices in that period to provide strong support for particular run estimates. When the harvest component of the model was removed from the calculation of the negative log-likelihood (i.e., "no-harvest model"), the model was much more stable across starting values in ADMB, though it still found two solutions (Figs. 8, 13, 14). When the new mark-recapture estimates were not used and over-dispersion parameters were not pooled by type (Fig. 13), removing the harvest component of the likelihood reduced both the mean (23,780) and maximum (75,090) difference among model solutions over both the base and pooled estimation models. When new mark-recapture estimates were used (Fig. 14), the no-harvest model improved the stability over the base model, but not the pooled estimation model, with a mean difference among solutions of 20,370, and a maximum difference of 81,410. Further reducing parameterization by pooling over-dispersion parameters, the pooled/no-harvest model was stable whether or not the new mark-recapture estimates were included (though it found different solutions between those two scenarios; mean differences among solutions across starting values less than 0.3 and maximum differences less than 3 for both scenarios; Fig. 15, 16). Thus, another key finding is that the only model

formulation that did not produce multiple solutions was one where the harvest component was excluded and where over-dispersion parameters were pooled into weir and aerial survey projects.

Starting-value Sensitivity in Estimation Model Using Simulated Data

In simulated data sets, the no-harvest model was stable (i.e., gave a unique solution in terms of annual estimates of run sizes) across starting values, regardless of whether over-dispersion parameters were pooled. This stability held across scenarios with synchronous sub-population dynamics (Figs. 17, 18), asynchronous sub-population dynamics (Figs. 19, 20), and with (Figs. 17, 19) and without (Fig. 18, 20) observation error. Thus, given that asynchrony and observation errors exist in the Kuskokwim Chinook system, much of the instability in ADFG's original run-reconstruction model appears to derive from an overly complex harvest component of the model that is not well parameterized.

The original estimation model tended to fit certain escapement indices more closely than others (Fig. 21), indicating that it was drawing the most information from these escapement data. The Kwethluk and Kogrukluq weir indices were very highly correlated (correlation > 0.95 across all estimation model structures) with their model estimated values, and Kogrukluq weir observations were also particularly correlated with the model estimated values. This pattern arose across all model structures (pooled/individual over-dispersion, with or without new mark-recapture estimates). In general, the estimation model better fit observed weir indices than aerial survey indices. The Bear River aerial index [labeled as 'a.ber' in Fig. 21] was particularly uninformative, as model estimates had low correlations (correlation = ~0.2) with observed escapement values (Fig. 21). Pooling over-dispersion parameters by escapement index type slightly reduced the correlation between observed and estimated escapement indices. Additionally, the R and ADMB implementations of the model fit the aerial indices differently, with the ADMB implementations fitting the Kwethluk aerial survey very well (correlations > 0.7), while the R model fit this index with a correlation of approximately 0.6. The R implementation also fit the Holokuk aerial survey ('a.hlk') more poorly (correlation ~ 0.2), while the ADMB implementation fit these observations better (correlation ~ 0.6). Interestingly, within each modeling implementation, the different model structures and data availability scenarios had little effect on the relative fits among streams (Fig. 21).

The original estimation model generally estimated higher over-dispersion parameters for weir indices than for aerial survey indices (Fig. 22). Higher over-dispersion parameters indicate less variation of observed escapement indices around the model predicted mean values, another measure of the amount of information the model derives from each index. This result is not surprising, as weir indices are much closer to census counts (though still only indices) than are aerial surveys, which are subject to variable flight conditions, viewing conditions, and non-linear bias across ranges of escapement (Jones et al. 1998). As such, weir indices should provide more consistent indicators of run size than aerial survey indices, though the value of information derived from weirs and aerial surveys should be quantified objectively, as discussed later in this report (see Discussion).

Simulation Model Experiments

While individual estimates within model iterations can have high errors, in a qualitative sense, the estimation model performed reasonably well across a range of simulated scenarios. The average relative error across model iterations in annual run estimates was always less than 15%, and typically less than 10% across the scenarios examined. The relative performance of the model declined as the operating model violated more assumptions of the estimation model, but even in scenarios violating several assumptions, the model still performed fairly well, and management targets set from these estimates tended to be biologically conservative and thus appear to not present basin-wide conservation problems (see *Management Implications of Estimation Errors* below).

It is important to note that the sensitivity of the estimation model to starting parameter values, as described above, was also seen in our simulation experiments. As such, we may have occasionally missed the global likelihood minima so that the errors and biases of true best model fits may be lower than what we present here. However, the relative impacts of violating different assumptions about the underlying population dynamics, or of different estimation model structures, should hold and provide information about which model assumptions are most important to address as future model development efforts progress.

Accuracy and Bias

During explorations of run-reconstruction model fits to simulated data, we identified a pattern wherein the model consistently overestimated total escapement and thus total run size. This issue stemmed from the pooling of harvest and effort data across the final three weeks of the data set. As the CPUE likelihood does not assume a linear relationship between catch and effort, this pooling across weeks in the original model is statistically inappropriate. Fitting the run-reconstruction model to actual data without pooling catch and effort across weeks resulted in lower run estimates for all years after 1990, regardless of which parameterization or mark-recapture data scenario we examined (Figs. 23, 24). The largest effect of the pooling of catch and effort data over the final three weeks of each season occurred when over-dispersion parameters were pooled, and when there were no new mark-recapture estimates included in the run-reconstruction (Fig. 25).

Asynchronous sub-population dynamics increased the normalized RMSE of annual run estimates across all scenarios and all estimation model structures when compared with synchronous cases (i.e., as assumed by the original ADFG estimation model; Fig. 26-28). This result was expected, because the estimation models all assume that each sub-population accounts for a constant proportion of the total run through time; asynchrony violates this assumption. Additionally, asynchronous sub-population dynamics increased the annual relative error in run estimates (from ~2.5 - 9% to ~4 - 15%; Fig. 29-34), as well as the bias across the time series (Fig. 35-40).

The presence of regime shifts generally increased the normalized RMSE in run estimates, though the effect was strongest when sub-population dynamics were asynchronous (Fig. 27). Annual relative error in run estimates also increased in the presence of regime shifts (Fig. 31, 32, 37, 38), particularly with asynchronous dynamics among sub-populations. Pooling over-dispersion parameters reduced the impacts of regime shifts in asynchronous sub-populations on the bias in run estimates (Fig. 37, 38). Pooling had much less of an effect when the sub-populations were synchronous.

Non-linearity in aerial survey escapement indices, as opposed to linearity in the original model, increased the normalized RMSE in run estimates, and the effect was particularly strong when sub-population dynamics were asynchronous (Fig. 28). Non-linear biases in aerial surveys generally increased the bias across the time series (compare Figs. 39, 40), as well as increased the average relative error in each individual year (compare Figs. 33, 34). The bias and magnitude of annual run estimate relative error caused by non-linear biases in aerial surveys were greatly diminished when new mark-recapture estimates were incorporated into the model and when over-dispersion parameters were pooled by index type.

Incorporating information from recent mark-recapture experiments reduced the normalized RMSE across all scenarios and model structures compared to omitting those recent mark-recapture data (Fig. 26-28). The benefit was smaller when sub-population dynamics were asynchronous, due to all mark-recapture events being towards the end of the time-series. Early years of the time-series were still subject to higher errors when sub-population dynamics are asynchronous. The effect of including new mark-recapture estimates on the average annual relative error (Fig. 29-34) and consistency of bias (Fig. 35-40) was striking. Including the new mark-recapture estimates greatly reduced the relative error in the last half of the time series (from ~12-15% to ~4-5%). The relative error in scenarios with asynchronous sub-population dynamics and new mark-recapture data were less than the relative error in scenarios with perfect synchrony in sub-population dynamic without new mark-recapture. Without these new mark-recapture estimates, relative error quickly increased after the initial set of mark-recapture estimates. These results highlight the importance of periodic mark-recapture studies for reducing relative error in run size estimates in the future, though they are not likely to improve the estimation accuracy in the early part of the time-series.

Removing the harvest component from the estimation model reduced the normalized RMSE of annual run estimates compared to the original estimation model (Fig. 26). However, removal of the harvest component tended to the average magnitude of relative error in annual estimates in the early portion of the time series (Fig. 29, 30), when there were limited escapement counts. Under these situations the harvest component likely provided much of the information used to estimate each annual run value. This result is unexpected and should be explored further.

While the model estimation errors were generally small in Figs. 29-34, the confidence intervals on those estimates did not always include the true run size value. The proportion of simulations in which model estimates included the true values was influenced by the biological and management scenario examined (Fig. 41). The presence of asynchronous sub-population dynamics generally reduced the coverage probability of model estimates relative to scenarios with synchronous sub-population dynamics. If the population dynamics were characterized by asynchrony among stock components, the coverage probability tended to be only about 60%.

The presence of regime shifts slightly reduced the coverage of model estimates when sub-population dynamics were asynchronous, especially in the early years, but had either little or the reverse impact on coverage when sub-population dynamics were perfectly synchronous. The incorporation of recent mark-recapture data modestly increased the coverage of model estimates for scenarios with asynchronous sub-population dynamics, but had no effect on coverage for scenarios with synchronous sub-population dynamics. Coverage values were generally highest in the earliest years with few escapement indices available, as these estimates had greater uncertainty and wider confidence intervals.

Management Implications of Estimation Errors

The consequences of inaccuracy and bias produced by the estimation model results depend on how those results are used for management of Kuskokwim River Chinook salmon. Because the run-reconstructions are used to estimate spawner-recruit relationships that inform the setting of escapement goals, we examined how the accuracy and bias issues identified above can affect management reference points (i.e., escapement goal targets). First, we examined whether the abundance estimates derived from the run-reconstruction model produced spawner-recruit relationships that indicated over-compensatory dynamics in the system (as is suggested by the existing data on Chinook salmon in the Kuskokwim River). We then examined the effect of assuming over-compensation on management reference targets by estimating both escapement goals and harvest rates that would produce MSY and comparing them to those values that produced MSY when populations were simulated forward in time across the range of potential harvest rates.

Catalano (2012) found that appearance of strong overcompensation in recruitment dynamics in Kuskokwim River Chinook salmon was anomalous when compared with the eleven other Alaska Chinook salmon stocks included in his analysis where there was no compelling support for overcompensation in population dynamics. Because productivity regime shifts have been suggested as a potential reason for the apparent strong over-compensation detected in the Kuskokwim River Chinook salmon population, we examined the effect of regime shifts of different strengths and durations on the likelihood of detecting overcompensation and the subsequent effects on management targets. We ran these scenarios with both synchronous and asynchronous sub-population dynamics, though all were produced by Beverton-Holt production functions. We assumed that all mark-recapture data were available for all scenarios.

The Shepherd spawner-recruitment model fit to the data from brood tables derived from the run-reconstruction model suggest that the likelihood of detecting over-compensatory dynamics (i.e., a Ricker shaped spawner-recruit relationship rather than a Beverton-Holt shaped relationship, as indicated by Shepherd's $\beta > 1$) was affected by the presence and characteristics of regime shifts acting upon the population. Stronger regime shifts produced the appearance of stronger overcompensation in recruitment dynamics (Fig. 42). Further, the duration of regimes played a major role in whether over-compensatory dynamics were detected; strong overcompensation was detected more consistently when regime duration was roughly equal to the longevity of Chinook salmon (4 to 8 years). The effects of regime strength and duration were present regardless of whether sub-population dynamics were synchronous. Asynchronous dynamics generally produced larger estimates of over-compensation strength (i.e., Shepherd's β) even if population dynamics were produced via Beverton-Holt recruitment.

We found that Ricker spawner-recruit relationships fit to the output from the run-reconstruction model generally identify S_{msy} values larger than those that would produce MSY based on long-term simulations, regardless of whether sub-population dynamics were synchronous or asynchronous (Fig. 44). While estimates of S_{msy} are high for both synchronous and asynchronous dynamics, misspecifications are smaller on average when dynamics are asynchronous. Similarly, estimates of the harvest rate that would produce MSY (as derived from the Ricker spawner-recruit relationship fit to estimated run sizes) underestimated the optimal harvest rate (Fig. 43). Thus, the presence and characteristics of periodic productivity regime shifts affected estimates of management targets in our simulation models. Estimation errors generally increased with regime shift strength and were strongest when the duration of each regime state roughly matched the average longevity of individuals in the populations being simulated.

From a conservation perspective, misspecifying the production dynamics (as driven by Ricker dynamics rather than by Beverton-Holt dynamics) in the stock-recruitment analysis produces management reference points that result in biologically conservative harvest rates. That is, more Chinook salmon are allowed to spawn than would produce the maximum sustainable yield at equilibrium. Such an approach will reduce the risk of overharvesting (or causing the extinction of) sub-stocks that are less abundant or productive than other sub-stocks at any point in time. Such contrasts in the productivity of sub-stocks should be expected in the Kuskokwim, given the existing data suggesting little coherence in the dynamics of sub-stocks in the system, and the prevalence of such biocomplexity in other salmon ecosystems. Even though the Kuskokwim River aggregate population has apparently supported high harvest rates while at low abundance in the past before rebounding to high abundance, small populations are inherently at greater conservation risks than large populations. Stochastic mortality events, Allee effects, and other compensatory mortality processes could all result in reduced productivity at low abundances. Thus, using a Ricker relationship to set management targets is a less risky approach for conserving stock diversity even if the underlying dynamics have no over-compensation in them.

From the perspective of harvest potential, using a Ricker model for management leads to lower harvest rates than those that would produce the maximum sustainable yield. These results suggest that this management approach could be reducing harvest opportunities, if the population is truly driven by Beverton-Holt dynamics subject to productivity regime shifts. For subsistence fishing, such overestimation of the optimal escapement goal increases the risk of unnecessary restrictions on harvest which may have important social consequences for communities reliant on Chinook salmon as a nutritional and cultural resource.

The trade-offs between providing harvest opportunity and conserving stock diversity under different management strategies and different levels (and types) of uncertainty should be rigorously explored to inform management about the most effective way to balance multiple objectives in a system with high levels of uncertainty concerning the dynamics of the ecosystem. Such management strategy evaluations are increasingly important as resources for research and monitoring become scarcer, and as the ecosystem is subjected to new perturbations such as those caused by ongoing climate change and by the poorly understood erosion of the age-structure of the population that has been observed in recent decades.

Comparison of Original and Revised Model on Stock-Recruit Analysis

We examined the relationship between spawning stock and recruitment inferred from output from the original run-reconstruction model compared to a revised model that eliminated some of the technical problems we identified with its structure. In particular, the revised model included log-normally distributed errors on data collected from weirs and aerial surveys, the variances were pooled for each of these survey types, and the harvest model was also reformulated. The harvest model was changed to an annual passage adjusted CPUE model, and the likelihood was changed to a log-normal likelihood with common variance across years and gear types according to:

$$CPUE_{jyg} = \sum_j \frac{H_{jyg}}{F_{jyg}}$$

$$P_{yg} = \sum_j p_j \text{ for all } j \text{ where } G_j = g$$

$$\sigma_q^2 = \ln(cv_q^2 + 1)$$

$$LL(q, \hat{N}_y, cv_q | H_y, F_y, p_y, G_y) = \sum_g \ln\left(\sqrt{\sigma_q^2}\right) + \frac{\left(\ln\left(\frac{CPUE_{yg}}{P_{yg}}\right) - \ln(q_g \hat{N}_y)\right)^2}{2\sigma_q^2}$$

where G_j is the gear type g fished in week j , y indicates year, H indicates fisheries harvest in numbers of fish, F indicates fishery effort, p is the proportion of the total run available to the fishery during the weeks the fishery was open, q is the catchability coefficient, \hat{N}_y is the estimated total abundance of Chinook salmon in year y , and cv_q is the estimated coefficient of variation for the catchability coefficient. For each year and gear type (i.e., mesh size), catch-per-unit-effort was added across weeks and the proportion of the total run that was available during fishing periods was calculated. The summed CPUE was divided by the proportion of total run available during the open fishing periods and compared to the predicted CPUE given by the product of the total run and the catchability coefficient for the specific gear type.

These changes to the original model seemed to remedy the model instability problems described earlier in the report such that the model converged on a single solution independent of starting values. We then compared the output from the original model using the old lower river expansions, and the new lower river expansions (ADFG, *unpublished data*) to this revised model (using the new lower river expansions).

While the revised model changed the estimates of recruitment produced from individual brood years, much of the underlying character of this relationship remained unchanged (Figure 45). In particular, the revised run-reconstruction model output still suggests a ‘Ricker’ stock-recruit relationship with considerable overcompensation (i.e., very high escapements lead to reduced recruitment), though the revised model output suggests weaker overcompensation than the original model. In terms of establishing management reference points, output from each of these three models produced similar estimates of the escapement that would produce maximum sustainable yield (S_{msy}) of about 66,000 – 73,000 fish (Table 2). The revised model estimate of 72,000 with new expansions is slightly higher than the estimate produced from the original model with new expansions (66,000 fish). The range of escapements that would produce at least 90% of MSY spans from about 42,000 to 105,000 for all three models (Table 2, Figure 45)¹. Thus, the revised model (like the original model) leads to estimates of S_{msy} comparable to the lowest escapements observed in the system. Whether such low escapements protect stock diversity and maximize harvest potential from the system remains highly uncertain. The principal difference between the revised model and the original model (both with the old expansion factors) is that the stock is not as productive as suggested by the original model. In particular, MSY for the revised model is estimated to be

¹ Note: these MSY based escapement ranges are for the purpose of model review and diagnostics and are not being recommended for adoption as a formal escapement goal range.

about 20% lower than for the original model, although these are both suggested to occur at similar escapements.

Table 2. Comparison of the spawner abundance (S_{msy}) expected to produce maximum sustainable yield (MSY) for three different run-reconstruction model scenarios for Kuskokwim River Chinook salmon. Estimates for the original model are given with old or new lower river expansions of the mark-recapture estimates. Estimates from the revised model are only given using the new expansion factors. The upper and lower bounds for the range of escapements expected to produce at least 90% of MSY are also given for each of the model scenarios.

| Model Scenario | S_{msy} | Spawners Producing 90% of MSY | |
|-----------------------------------|-----------|----------------------------------|---------|
| | | Low | High |
| Revised Model, New Expansions | 71,911 | 46,138 | 102,335 |
| Original Model, New Expansions | 66,335 | 42,373 | 94,829 |
| Original Model, Old Expansions | 73,431 | 46,770 | 105,291 |

As described in the sections above, these analyses of the stock-recruitment relationship assume that the relationship between the environment and the productivity of the stock is not changing in any systematic way through time (other than year-to-year variation in productivity). Detailed analyses of this and other Pacific salmon stocks (e.g., Adkison et al. 1996; Peterman et al. 1998, 2003; Pyper et al. 2005; Dorner et al. 2008, 2018; Ohlberger et al. 2016; Peterman and Dorner 2012; Kilduff et al. 2014; Malick and Cox 2016) suggest several reasons for expecting that this relationship is not stationary through time. Climate-driven changes in the environment through such phenomena as the Pacific Decadal Oscillation (PDO) and the North Pacific Gyre Oscillation (NPGO) may cause low-frequency changes in stock productivity that affect the relationship between stock and recruitment. Evidence of long-term changes in the age structure of Kuskokwim River Chinook salmon, characterized by the disappearance of the oldest and largest fish from the population (Lewis et al 2015, Ohlberger et al. 2018), suggests that other poorly understood shifts in the ecology of Chinook salmon, potentially combined with selective harvest, may be producing these trends. The consequences of these changes in demographic structure for the quality and egg mass of spawners (and therefore the stock productivity) are largely unknown and are currently being investigated by a separate AYK SSI Chinook Salmon Escapement Quality Expert Panel. However, the direction and magnitude of change is cause for concern, as demographic shifts driving declines in escapement quality across the time-series can introduce bias into run-reconstruction and spawner-recruit model results. Therefore, we encourage vigilance in watching for changes in the productivity of the Kuskokwim River Chinook stocks by maintaining appropriate fishery and escapement monitoring. Though it is beyond the

scope of this model review, we recommend re-evaluation of the stock-recruit relationship in the near future (and routinely thereafter) with explicit consideration for the potential effect of declining escapement quality over the time series (e.g., Jones et al (2018) recent analysis of the effect of changes in escapement quality on spawner-recruit model for Canadian-origin Chinook Salmon in the Yukon River).

Further, given the increasing recognition of the prevalence of stock structure within the Kuskokwim River Chinook salmon stock complex, it is critical to assess whether the current stock-recruitment analysis (which assumes a single homogeneous stock) produces management reference points that provide robust measures for protecting sub-stocks that may experience periods of low productivity but are harvested simultaneously with larger or more productive stocks. Thus, we also recommend more detailed assessment of the stock-recruit dynamics of the overall system with an eye towards understanding (1) the consequences of different harvest strategies for maintaining biocomplexity within the stock complex, and (2) the long-term consequences for sustainability of the stocks under future unknown environmental shifts.

Recommendations

Based on our review and analysis of the original stock-assessment models used by ADFG for managing Chinook salmon in the Kuskokwim River, we offer the follow list of recommendations for improving the performance of these models and, thus, scientifically-based management of this invaluable resource.

Run-Reconstruction Model Structure

- 1) Pooling the over-dispersion parameters in the estimation model reduced the normalized RMSE and bias in run estimates. Additionally, this improved the model stability across starting parameters, particularly when paired with mark-recapture estimates across multiple periods. Pooling the over-dispersion parameters is easily implemented and should improve accuracy and reduce bias. However, pooling over-dispersion parameters alone was not sufficient to fully correct the estimation problem in which there were multiple solutions.
- 2) The original model assumes that the errors on weir and aerial survey data are distributed according to a negative binomial distribution. We recommend changing the model to more appropriately assume that these errors are log-normally distributed. While we have not included detailed analysis in this report, analysis of simulated data showed that changing the model to include this error structure improved its stability, albeit modestly. We can provide a more thorough summary of results upon request, but have elected to not include them in this report.
- 3) Removing the harvest component of the likelihood from the estimation model greatly improved model stability across starting values, which was unexpected. The lack of estimates of total in-river abundance, particularly in the early portion of the time-series, coupled with asynchronous dynamics among sub-populations, generated multiple local minima on the likelihood surface, leading to critical minimization problems. However, removing the harvest component of the likelihood also increased the bias in estimates during the early part of the time-series (as shown in simulation experiments), because run-estimates in early time-series no longer derived any information from the harvest data and relied solely on relatively few escapement indices. Thus, there are trade-offs between model

stability and estimation bias when determining the value of retaining the harvest component of the likelihood. The harvest component of the original model appears to be either misspecified or overly complex, and we recommend that this part of the model be scrutinized and alternative formulations considered as a way to improve model performance.

- 4) The original estimation model is highly sensitive to parameter starting values, particularly when fit in ADMB (though ADMB found the solutions with the best fit by negative log-likelihood). Even the revised model structures are sensitive to starting values, though less so than the original model. As such, the model should be fit across a range of starting values for parameters to explore the likelihood surface. Negative log-likelihoods of the different solutions should be compared across the range of solutions to ensure the best possible fit is being found. As described above, this poor convergence behavior suggests that the model is overly complex given the information content of the data it is fit to.
- 5) Any new model should be thoroughly tested using simulation experiments. All models are simplified representations of reality and, thus, require many assumptions about the nature of real-world dynamics. The potential impacts of violating these assumptions to management activities should be thoroughly investigated using simulation where the true underlying state of the system is known and biases or inaccuracies arising from assumptions in the estimation model can be determined.

Data Collection

- 1) Mark-recapture estimates of river-wide abundance are critical for tracking long-term changes in the stock, especially given that asynchrony and regime shifts are involved in the production dynamics of Chinook salmon in the Kuskokwim system. These are the only sources of information that anchor the magnitude of annual runs, allowing the estimation of all other parameters in the model. At least some level of sub-population asynchrony should be expected among the Chinook salmon populations of the Kuskokwim River, and without future mark-recapture estimates, run estimation errors will become continually less accurate and more biased as the relative production of different populations changes. Periodic mark-recapture data are extremely valuable for capturing the variation in population productivity that can occur in the Kuskokwim River. These data need not be collected each year, but if data are collected periodically, the model will be able to adjust its predictions around the time-varying productivity of individual populations. Thus, while funding for such projects may be limited at present, capitalizing on periodic funding opportunities that may arise is likely to be highly valuable.
- 2) How often mark-recapture programs should be conducted should be determined using a formal, quantitative value-of-information (VOI) approach based on simulations similar to those used by this panel but with added economic variables. Such VOI analyses will determine how the large expenses of mark-recapture programs compare with the value of the additional catches (and/or other benefits not quantifiable in dollar terms) taken in future years, and how often such programs are warranted for reducing model uncertainties. Link and Peterman (1998) conducted this type of VOI analysis to determine whether a fish wheel for sockeye salmon on the Nass River in British Columbia was worthwhile; it was. This same VOI approach could also be used to determine the value of various weir and aerial survey programs, which are at risk of becoming discontinued due to budget constraints.

Discontinuing such assessment programs without an objective assessment of the costs and benefits (i.e., in being able to properly assess the stock status) would be careless.

- 3) More explicit examination of the trade-offs between maintaining a collection of weirs and aerial surveys versus investing research funding in whole-system mark-recapture estimates of abundance should be done to optimize future research funding investments. We suspect that maintaining a network of weirs and aerial surveys, or investment in other ways to estimate sub-stock dynamics (e.g., through otolith microchemistry or genetic sampling) across the watershed will be critical for producing reliable abundance estimates in the future. In situations where there is considerable asynchrony among sub-stock dynamics, the value of any individual weir is likely to increase markedly as the number of sites monitored declines. However, these trade-offs have not yet been quantified and should be made a high priority for future research.
- 4) Systematic and non-linear bias (underestimation of abundance at high actual abundance) is well-documented in aerial surveys of spawning salmon. As the number of spawners in a system increases, the magnitude of the bias increases. This creates a problem with using aerial indices in an estimation model that assumes constant proportionality between indices and spawner abundance. This situation means that, in comparison, the escapement data provided by the weir projects are particularly valuable during years when no mark-recapture studies are conducted. The original estimation model generally fits higher over-dispersion parameters for weir projects than aerial surveys, indicating lower variance of observed weir escapement relative to the model's expected weir escapement. When decisions about funding of future monitoring programs are being made, the presumed greater value of information from weir projects relative to aerial surveys should be considered. Here again, formal value-of-information analyses would be extremely informative for those funding decisions.
- 5) Recent analyses of patterns in productivity of various Chinook salmon stocks suggest that they have become more coherent (positively correlated) in recent years (Kilduff et al. 2014; Dorner et al. 2018), as has been observed for sockeye salmon (Peterman and Dorner 2012) and pink and chum salmon (Mallick and Cox 2016). Such changes in the degree of synchrony among stock components is one expression of non-stationarity in population dynamics, which may further affect performance of the run-reconstruction model if Kuskokwim Chinook salmon are characterized by similar changes in population dynamics within the stock. A simulation approach similar to what we have used here could be used to explore the consequences of such changes in the ecosystem, though we anticipate that problems arising from changes in synchrony will probably be minor compared to changes driven by system-wide regime shifts in productivity, or to situations where data are sparse for estimating model parameters.

Stock-Recruit Analysis

- 1) The relationship between spawning stock and subsequent recruitment forms the biological basis for estimating the escapement that will maximize sustained production from salmon-producing ecosystems. In simplified form, such as that currently used to assess the productivity of Chinook salmon in the Kuskokwim River, it is assumed that the per capita reproductive potential of the population has remained constant over the course of the observed time series. However, given the

widespread observation that the size- and age-structure of the population has shifted towards smaller and younger individuals through time, this assumption is clearly tenuous, as the escapement quality may have deteriorated through time in response to currently unknown reasons. However, given the uncertainties in the stock assessment process for Kuskokwim Chinook salmon, it remains unclear whether the productivity of the stock and the management reference points derived from the stock assessments are changing in response to shifting age- and size-structure. We recommend re-evaluation of the stock-recruit relationship with a) explicit consideration for the potential effects of declining escapement quality (i.e., lower per capita reproductive potential) over the time series, and b) evaluation of potential harvest strategies focused on maintaining the largest, oldest segment of the population that are disproportionately female, as has been applied to the Kenai River Chinook salmon stocks in recent years (Fleischman and Reimer 2017).

- 2) Current stock assessments of Kuskokwim Chinook salmon assume that it is composed of a homogeneous stock distributed across the watershed, with each tributary contributing a constant proportion of the total run at any point in time. What is more likely, given accumulating knowledge in this and other systems, is that it is actually a stock complex composed of many populations (sub-stocks) that show some degree of demographic independence. What is not known is how the current management approach of assuming a single large stock for the purpose of setting allowable harvest rates risks eroding the diversity of populations that compose the stock complex. Harvest rates established by assuming a homogeneous stock run the risk of overexploiting unproductive sub-stocks. Thus, we recommend more detailed assessment of the stock-recruit dynamics of the overall system to assess: a) the consequences of different harvest strategies for maintaining the integrity of sub-stocks (i.e., the biocomplexity) across the entire watershed, and 2) the long-term consequences for sustainability of the stocks under future known environmental conditions. Trade-offs between risk to biocomplexity and harvest opportunity, as a function of the escapement goal, should be quantified and used in establishing and adapting escapement goals as new observations about stock dynamics accumulate.
- 3) The stock assessment process for Chinook salmon stocks in the Kuskokwim River remains challenging owing to substantial uncertainties in the data and in the ecological processes that generate variation in abundance. Formal incorporation of these uncertainties into the assessment process has improved through the development of a state-space approach for estimating the stock-recruitment relationship. However, whether uncertainties in the run-reconstruction analyses are properly integrated into the stock-recruitment analysis (that is used to inform management reference points) is not clear at present. The current practice of doubling the CV of the run-reconstruction estimates of abundance as input to the state-space analysis of stock-and-recruitment is arbitrary and may have unintended consequences on the interpretation of the biological status of the stocks. We recommend more formal exploration of the run-reconstruction model, using a simulation approach like we have adopted in this report, to develop a better understanding of the true uncertainties of the run-reconstruction estimates of annual abundance.

References

- Adkison, M. D., R.M. Peterman, M.F. Lapointe, D.M. Gillis, and J. Korman. 1996. Alternative models of climatic effects on sockeye salmon productivity in Bristol Bay, Alaska, and the Fraser River, British Columbia. *Fisheries Oceanography* 5:137-152.
- Alaska Department of Fish & Game. 2013. Chinook salmon stock assessment and research plan, 2013. ADFG Special Publication No. 13-01. Anchorage, Alaska.
- Bue, B.G., K.L. Schaberg, Z.W. Liller, and D.B. Molyneaux. 2012. Estimates of historic run and escapement for the Chinook salmon stock returning to the Kuskokwim River, 1976-2011. *Alaska Department of Fish and Game Fishery Data Series No. 12-49*.
- Catalano, M.J. 2012. A Comparative Analysis of Patterns in Productivity for AYK and Selected Non-AYK Chinook Salmon Stocks. Workshop Background Paper (unpublished working paper). In AYK SSI Chinook Salmon Synthesis Workshop, May 2-3, 2012. Arctic-Yukon- Kuskokwim Sustainable Salmon Initiative, Anchorage, AK.
- Conitz, J.M., K.G. Howard, and M.J. Evenson. 2015. Escapement goal recommendations for select Arctic-Yukon-Kuskokwim Region salmon stocks, 2016. *Alaska Department of Fish and Game Fishery Manuscript Series No. 15-08*.
- Dorner, B., M. Catalano, and R.M. Peterman. 2018. Spatial and temporal patterns of covariation in productivity of Chinook salmon populations of the Northeastern Pacific. *Canadian Journal of Fisheries and Aquatic Sciences* 75:1082-1095. <https://doi.org/10.1139/cjfas-2017-0197>.
- 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.
- Fleischman, S.J., and A.M. Reimer. 2017. Spawner-recruit analyses and escapement goal recommendations for Kenai River Chinook salmon. Alaska Department of Fish and Game, Fishery Manuscript Series No. 17-02, Anchorage, AK.
- Hamazaki, T., M. Everson, S.J. Fleischman, and K.L. Schaberg. 2012. Spawner-recruit analysis and escapement goal recommendation for Chinook salmon in the Kuskokwim River drainage. *Alaska Department of Fish and Game Fishery Manuscript Series No. 12-08*.
- Hamazaki, T., and Z. Liller. 2015. ADFG Memorandum: 2014 Kuskokwim River Chinook salmon run-reconstruction and model revisions. March 10, 2015.
- Hilborn, R. and Walters, C.J., 1992. Quantitative fisheries stock assessment and management. *Chapman-Hall: New York*.
- Jones, E.L., Quinn, T.J. and Van Alen, B.W., 1998. Observer accuracy and precision in aerial and foot survey counts of pink salmon in a southeast Alaska stream. *North American Journal of Fisheries Management*, 18(4), pp.832-846.
- Jones, M.L., S. Truesdell, J. Syslo, M.J. Catalano, and S. Fleischman. 2018. Developing tools to evaluate management strategies for sustainable exploitation of Yukon River Chinook Salmon. AYK-SSI Project Completion Report. 45 pp.

- Kilduff, D. P., L. W. Botsford, and S. L. H. Teo. 2014. Spatial and temporal covariability in early ocean survival of Chinook salmon (*Oncorhynchus tshawytscha*) along the west coast of North America. *ICES Journal of Marine Science* 71:1671-1682.
- Lewis B, Grant WS, Brenner RE, Hamazaki T (2015) Changes in Size and Age of Chinook Salmon *Oncorhynchus tshawytscha* Returning to Alaska. *PLoS ONE* 10(6): e0130184. doi:10.1371/journal.pone.0130184. pmid:26090990
- Liller, Z.W. 2017. 2016 Kuskokwim River Chinook salmon run-reconstruction and 2017 forecast. *Alaska Department of Fish and Game Regional Information Report 3A.17-02*.
- Link, M.R. and R.M. Peterman. 1998. Estimating the value of in-season estimates of abundance of sockeye salmon. *Canadian Journal of Fisheries and Aquatic Sciences* 55:1408-1418.
- Malick, M.J., S.P. Cox, F.J. Mueter, B. Dorner, and R.M. Peterman. 2017. Effects of the North Pacific Current on productivity of 163 Pacific salmon stocks. *Fisheries Oceanography* 26(3):268-281. doi:10.1111/fog.12190.
- 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.
- Mueter, F.J., J.L. Boldt, B.A. Megrey, and R.M. Peterman. 2007. Recruitment and survival of Northeast Pacific Ocean fish stocks: temporal trends, covariation, and regime shifts. *Canadian Journal of Fisheries and Aquatic Sciences* 64(6):911-927.
- Ohlberger, J., Scheuerell, M.D. and Schindler, D.E., 2016. Population coherence and environmental impacts across spatial scales: a case study of Chinook salmon. *Ecosphere*, 7(4).
- Ohlberger, J., E.J. Ward, D.E. Schindler, B. Lewis. 2018. Demographic changes in Chinook salmon across the Northeast Pacific Ocean. *Fish and Fisheries* oo:1-14, DOI: 10.1111/faf.12272
- Peterman, R.M., B.J. Pyper, M.F. Lapointe, M.D. Adkison, and C.J. Walters. 1998. Patterns of covariation in survival rates of British Columbian and Alaskan sockeye salmon stocks. *Canadian Journal of Fisheries and Aquatic Sciences* 55:2503-2517.
- 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., and B. Dorner. 2012. A widespread decrease in productivity of sockeye salmon populations in western North America. *Canadian Journal of Fisheries and Aquatic Sciences* 69 (8): 1255–1260.
- 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.
- Rogers, L.A. and Schindler, D.E., 2008. Asynchrony in population dynamics of sockeye salmon in southwest Alaska. *Oikos*, 117(10), pp.1578-1586.
- Schindler, D.E., Hilborn, R., Chasco, B., Boatright, C.P., Quinn, T.P., Rogers, L.A. and Webster, M.S., 2010. Population diversity and the portfolio effect in an exploited species. *Nature*, 465(7298), p.609.

Schindler, D., C. Krueger, P. Bisson, M. Bradford, B. Clark, J. Conitz, K. Howard, M. Jones, J. Murphy, K. Myers, M. Scheuerell, E. Volk, and J. Winton. Arctic-Yukon-Kuskokwim Chinook Salmon Research Action Plan: Evidence of Decline of Chinook Salmon Populations and Recommendations for Future Research. Prepared for the AYK Sustainable Salmon Initiative (Anchorage, AK). v + 70 pp.

Sharma, R., L. A. Vélez-Espino, A. C. Wertheimer, N. Mantua, and R. C. Francis. 2013. Relating spatial and temporal scales of climate and ocean variability to survival of Pacific Northwest Chinook salmon (*Oncorhynchus tshawytscha*). *Fish. Oceanog.* 22(1)14-31.

Figures

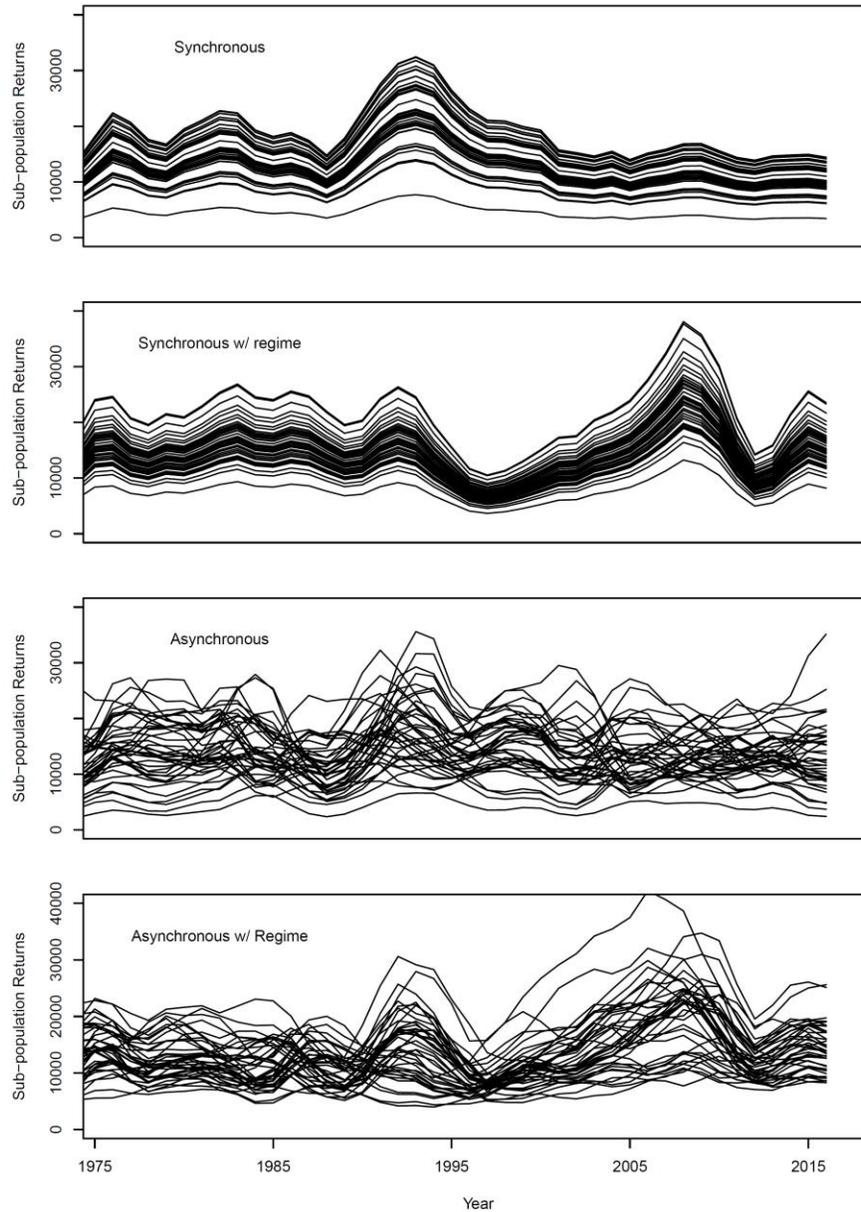


Figure 1. Example sub-population dynamics generated by the operating model under conditions with perfect synchrony among sub-populations (top), synchrony among sub-stocks but with regime shifts in productivity (top middle), conditions with asynchronous sub-population dynamics (bottom middle), and conditions with asynchrony among sub-populations and regime shifts in productivity (bottom). Each line represents the run size to a single sub-population.

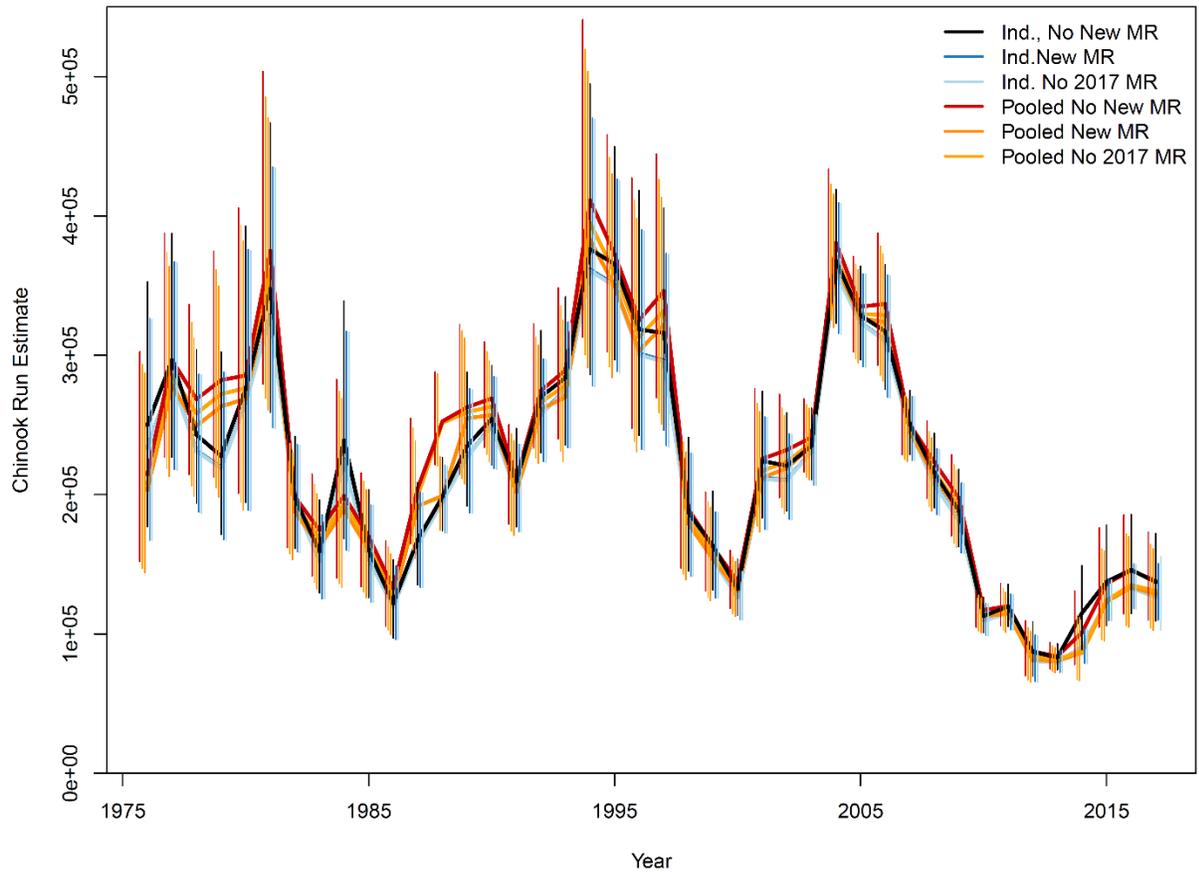


Figure 2. Run-reconstructions for Chinook salmon in the Kuskokwim River from the original model fit to observed data under different model structures (individual "Ind" or pooled over-dispersion parameters) and with different amounts of mark-recapture (MR) data available (no new mark recapture estimates, all new mark-recapture estimates, or new mark-recapture estimates but without the 2017 estimate). Starting values for these simulations are held constant, at the values currently used by ADFG, across all scenarios considered here.

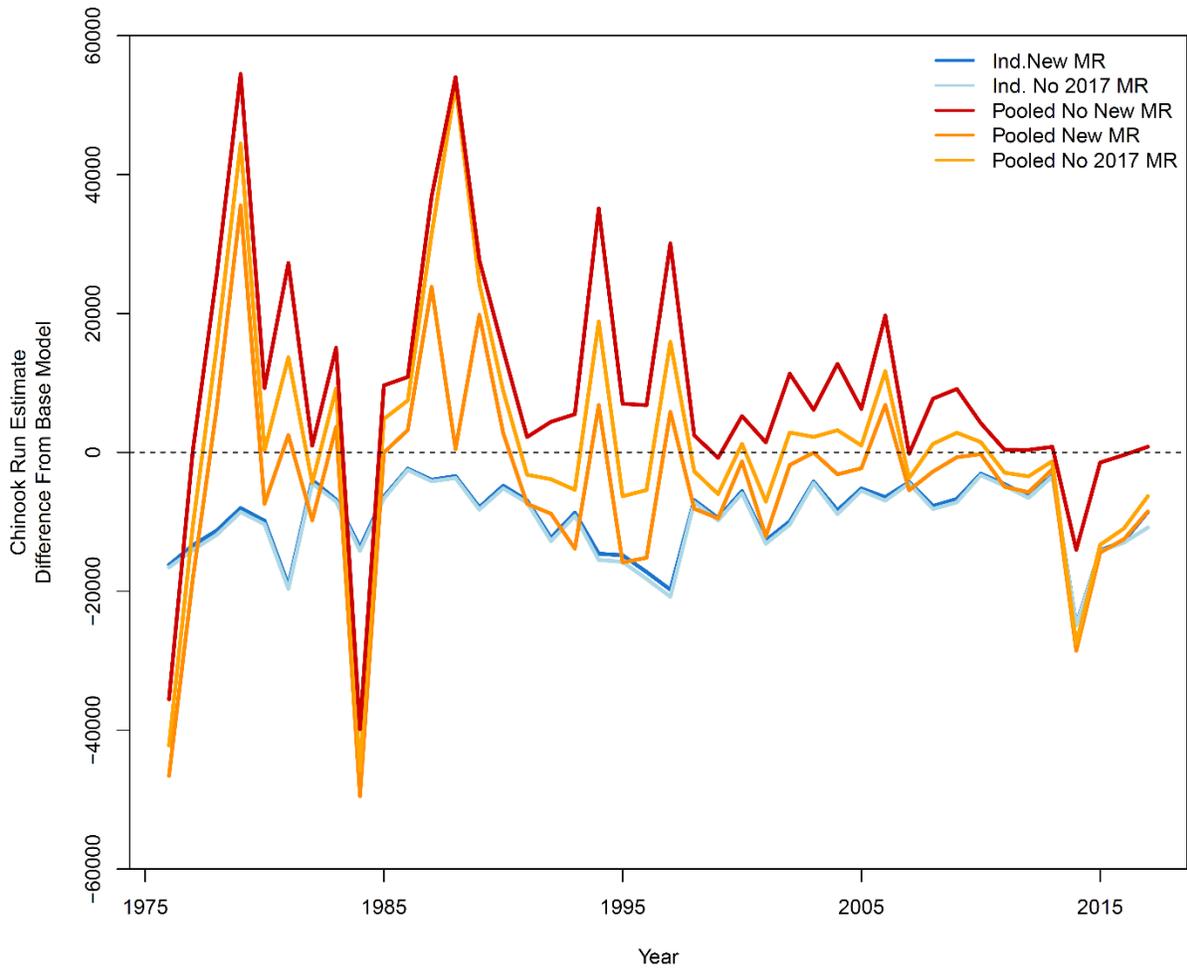


Figure 3. Changes in annual run-reconstruction estimates when switching from the original model (individual ("Ind") over-dispersion parameters and no new mark-recapture (MR) estimates available) to different model structures and data scenarios. Starting values for these simulations were held constant across all scenarios considered here at the values currently used by ADFG.

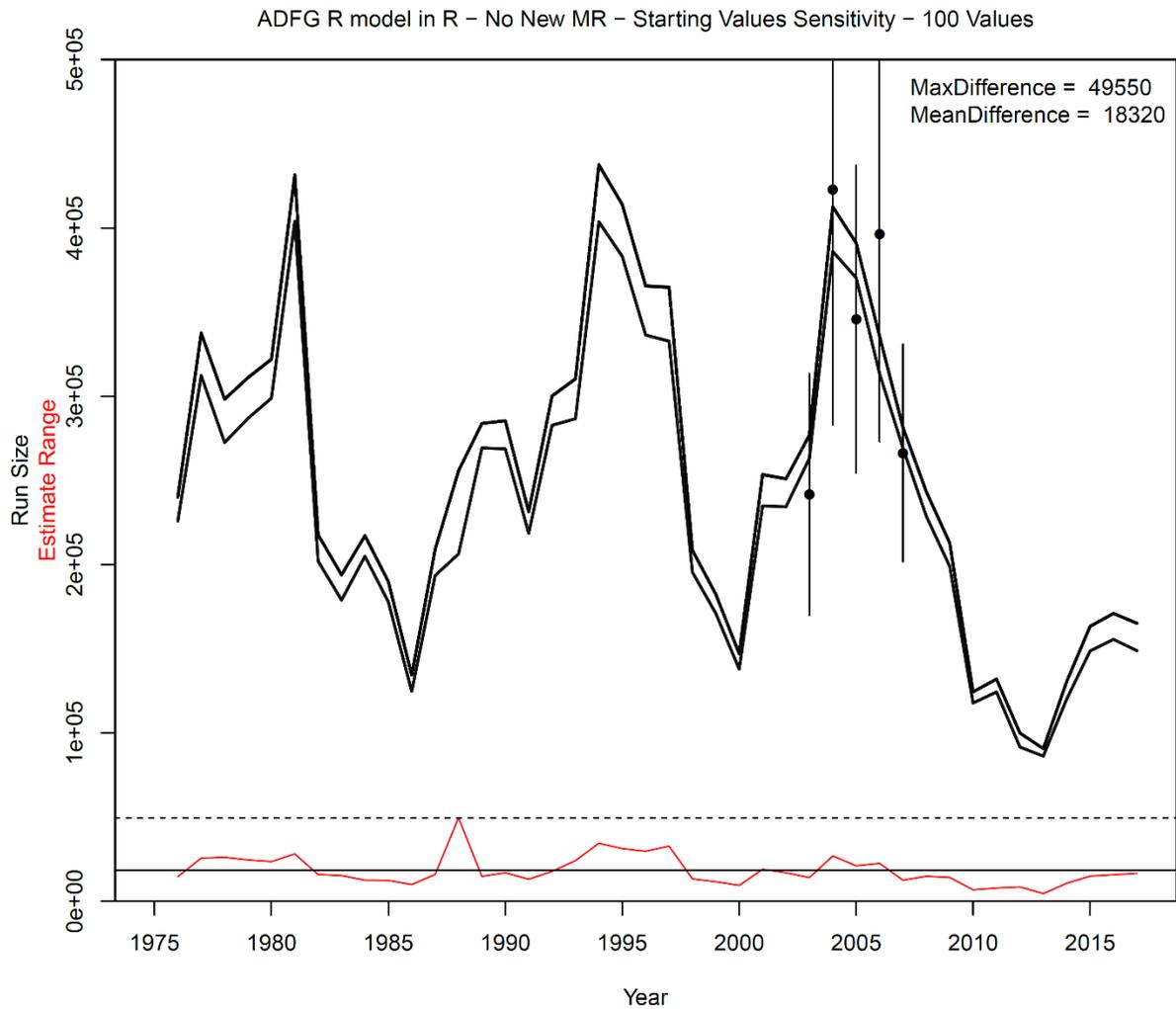


Figure 4. Run estimates across a range of starting values from the original ADFG R model implemented in R without new mark-recapture estimates. Black lines indicate stacked grey lines, representing repeated model convergence on some values. The red line indicates the range in model estimates for annual run size. The horizontal solid black line represents the mean of the estimate range, while the dashed horizontal line shows the maximum range across simulations. The dots with vertical bars are the means and 95% confidence intervals of the mark-recapture estimates.

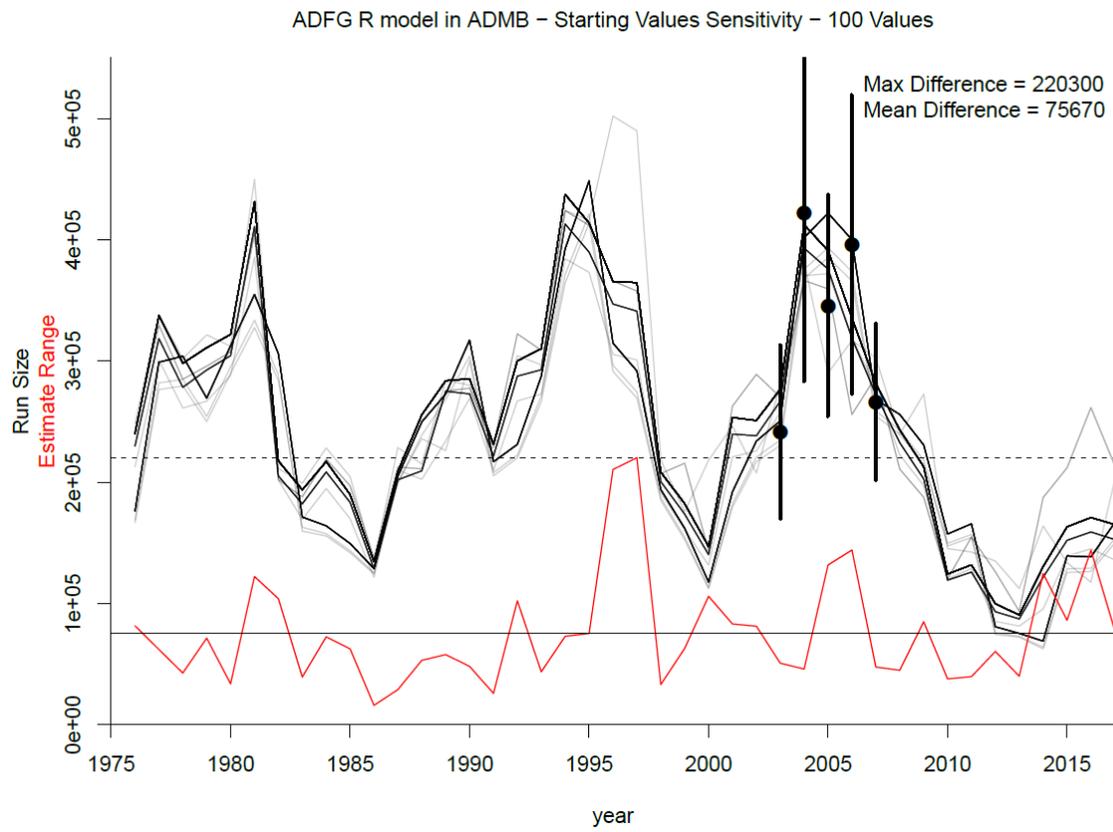


Figure 5. As in Fig. 4, but with the original run-reconstruction model implemented in ADMB.

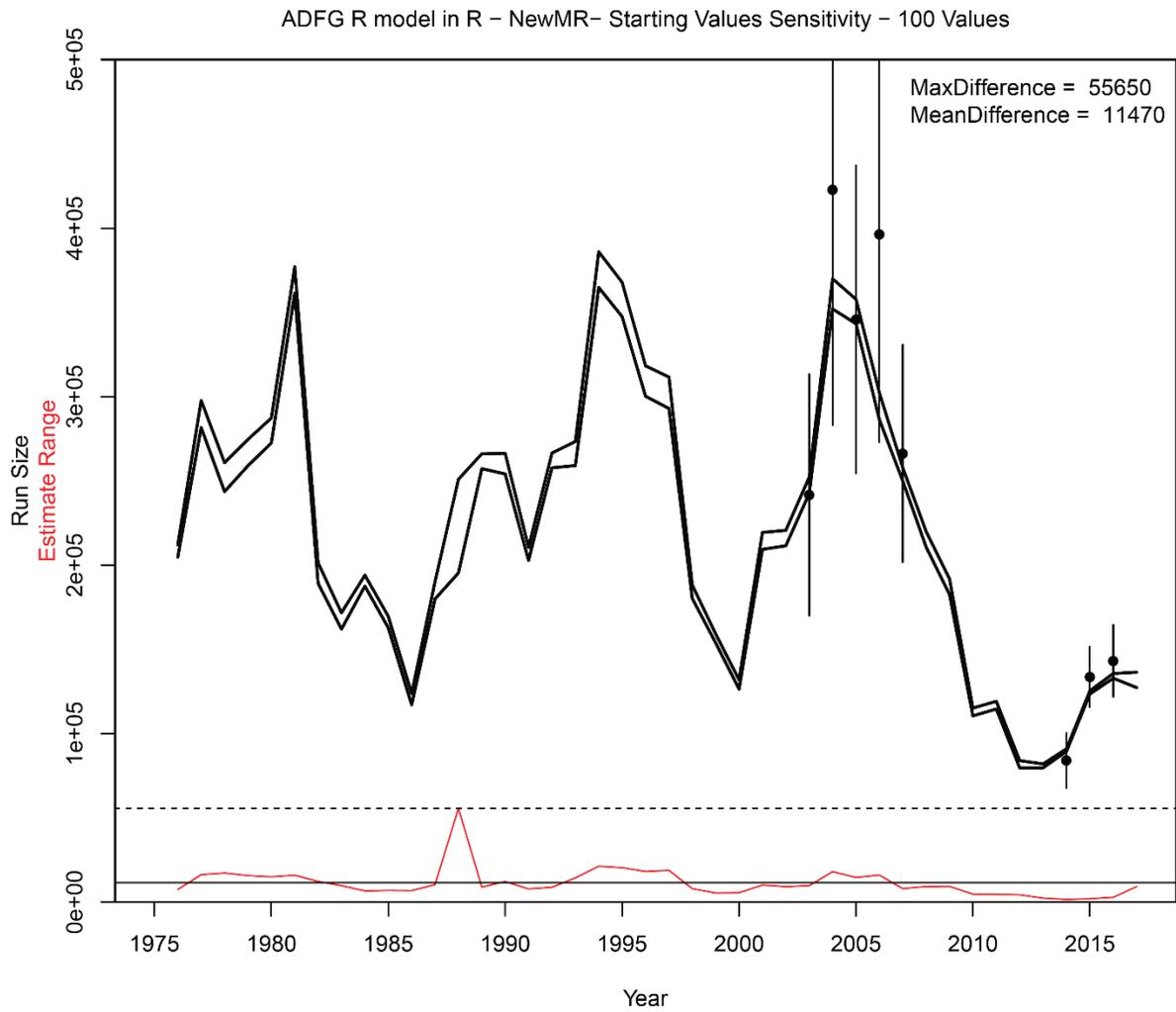


Figure 6. As in Fig. 4, but with new mark-recapture estimates available (2014-2016).

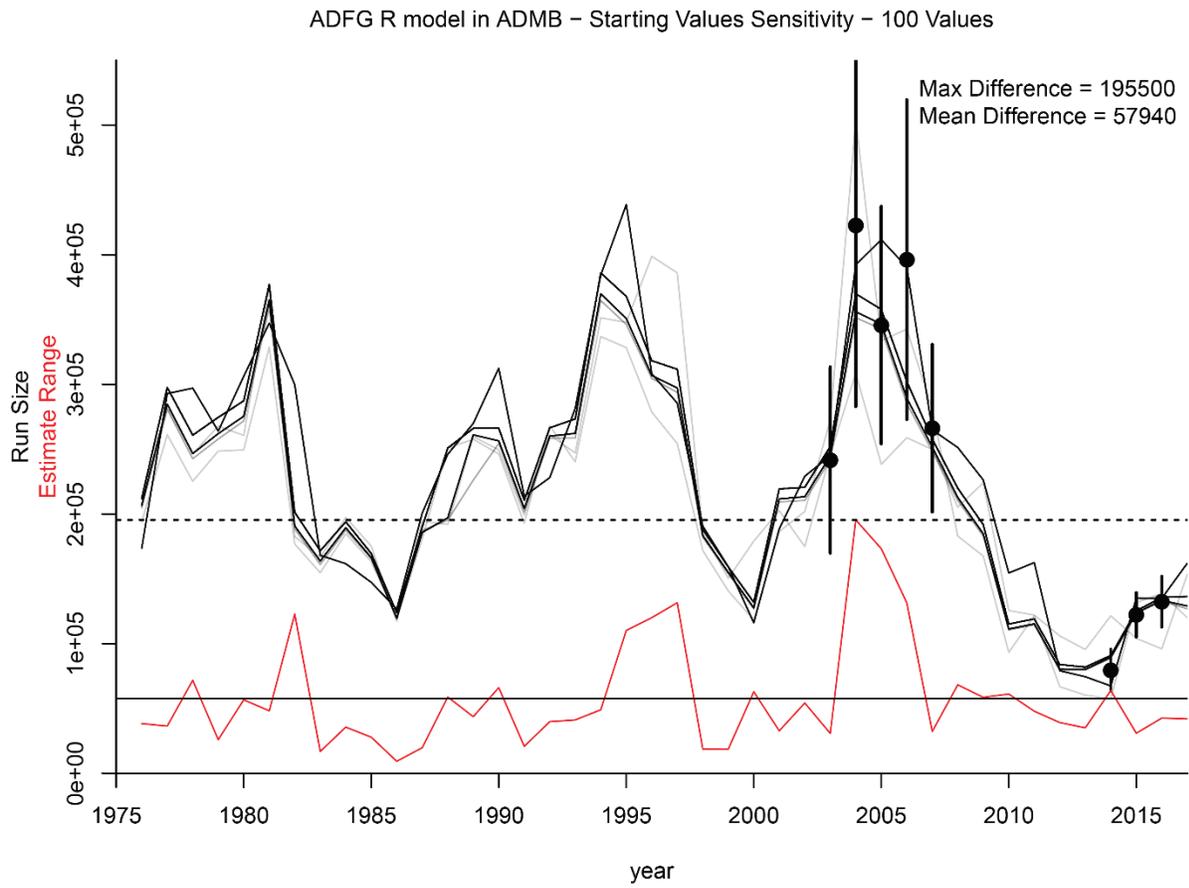


Figure 7. As in Fig. 5, but with new mark-recapture estimates available (2014-2016).

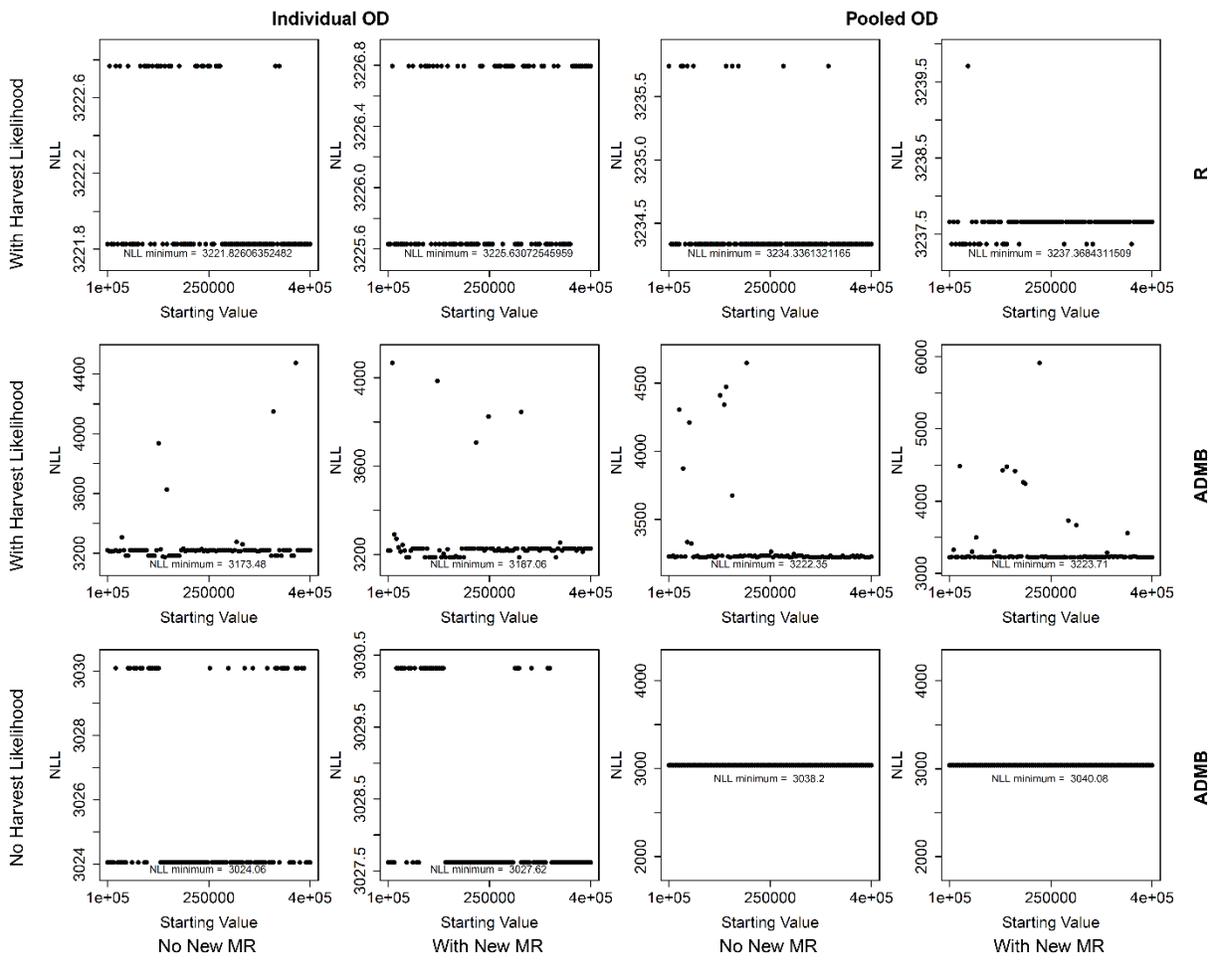


Figure 8. Negative log-likelihood values of original run-reconstruction model solutions across a range of starting values for annual run size. The top row show results for models fit in R, while the bottom two rows show results of models fit in ADMB. The top two rows show the results of models with the harvest component of the likelihood included, while the bottom row shows models fit without the harvest component of the likelihood. The left two panels are model fits with individual over-dispersion parameters for each escapement index, while the right two panels show model fits with over-dispersion parameters pooled by escapement index type. The first and third columns show model fits when no new mark-recapture data are available, while the second and fourth columns show model fits with new mark-recapture data available.

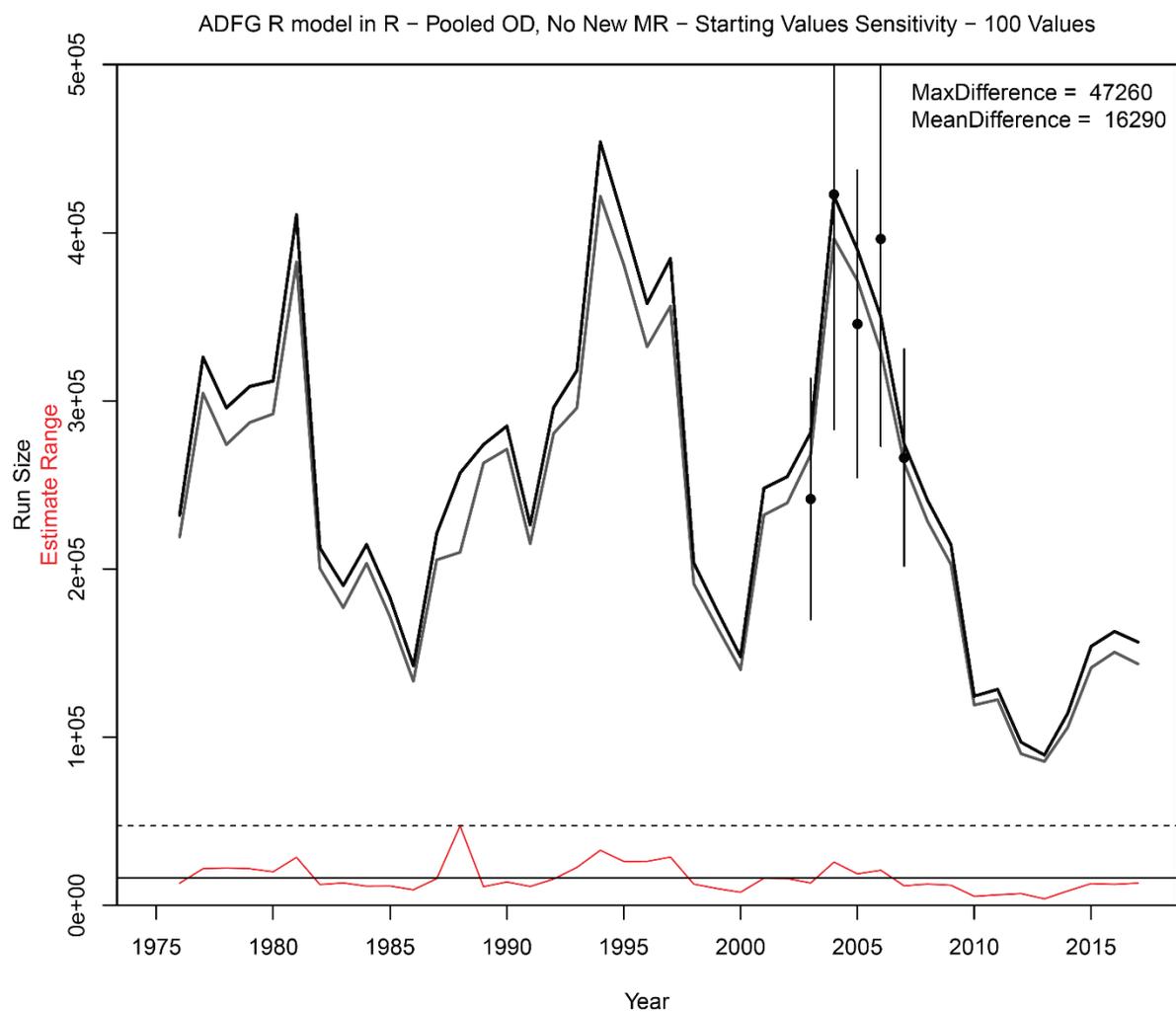


Figure 9. As in Fig. 4, but with over-dispersion parameters pooled by escapement index type (weir/aerial).

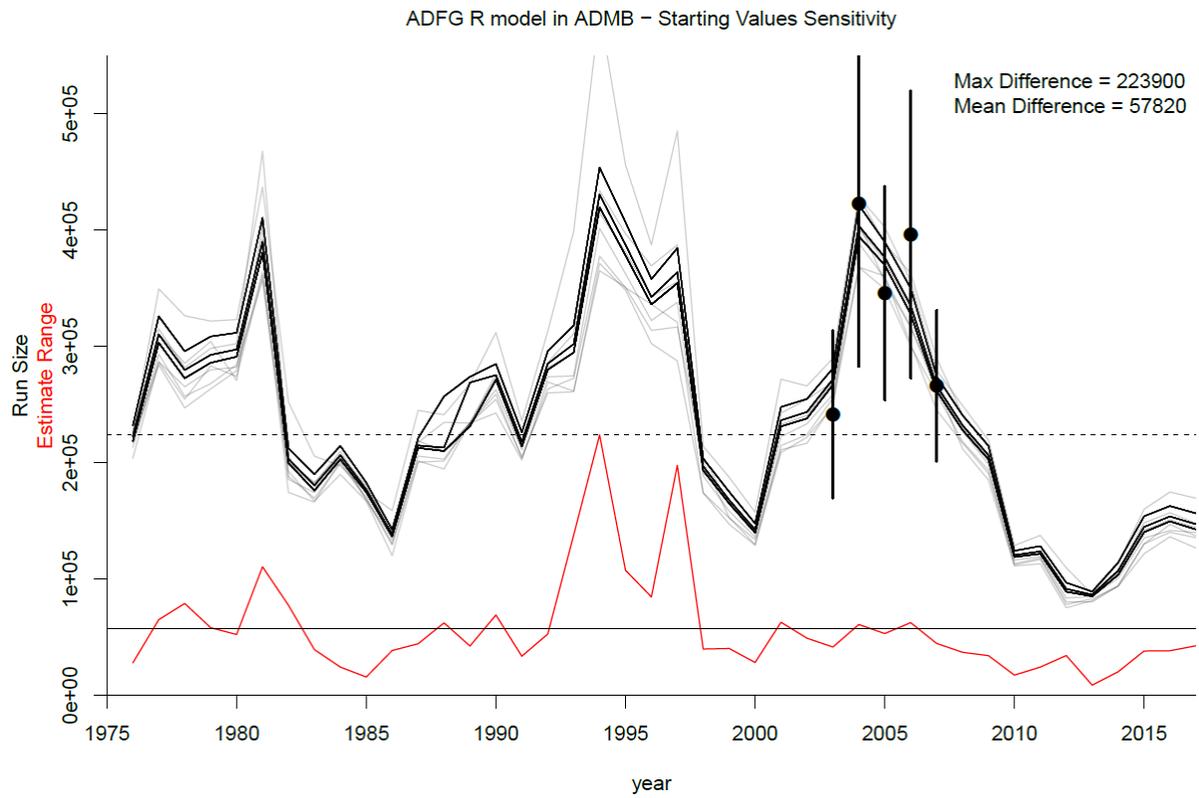


Figure 10. As in Fig. 5, but with over-dispersion parameters pooled by escapement index type (weir/aerial).

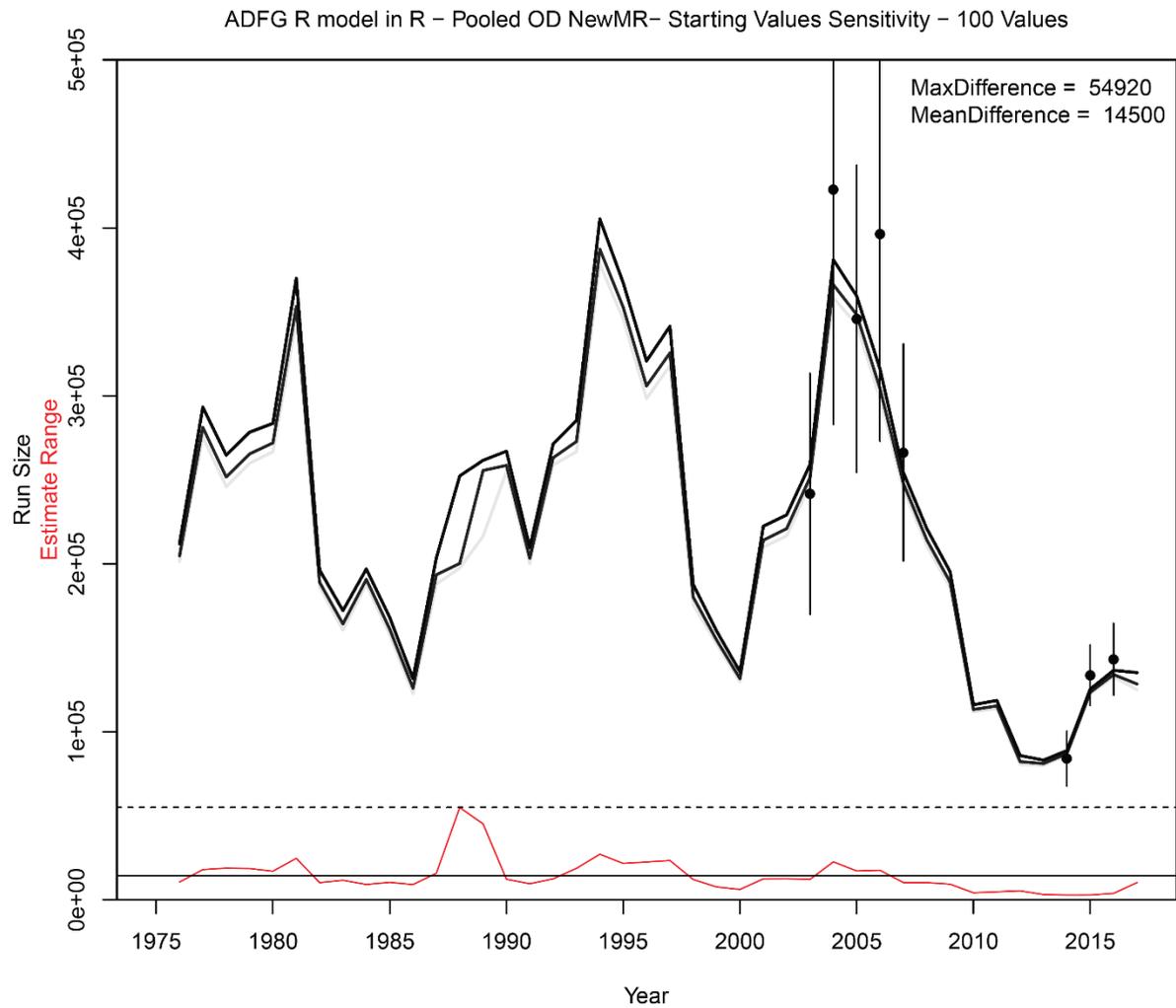


Figure 11. As in Fig. 6, but with over-dispersion parameters pooled by escapement index type (weir/aerial).

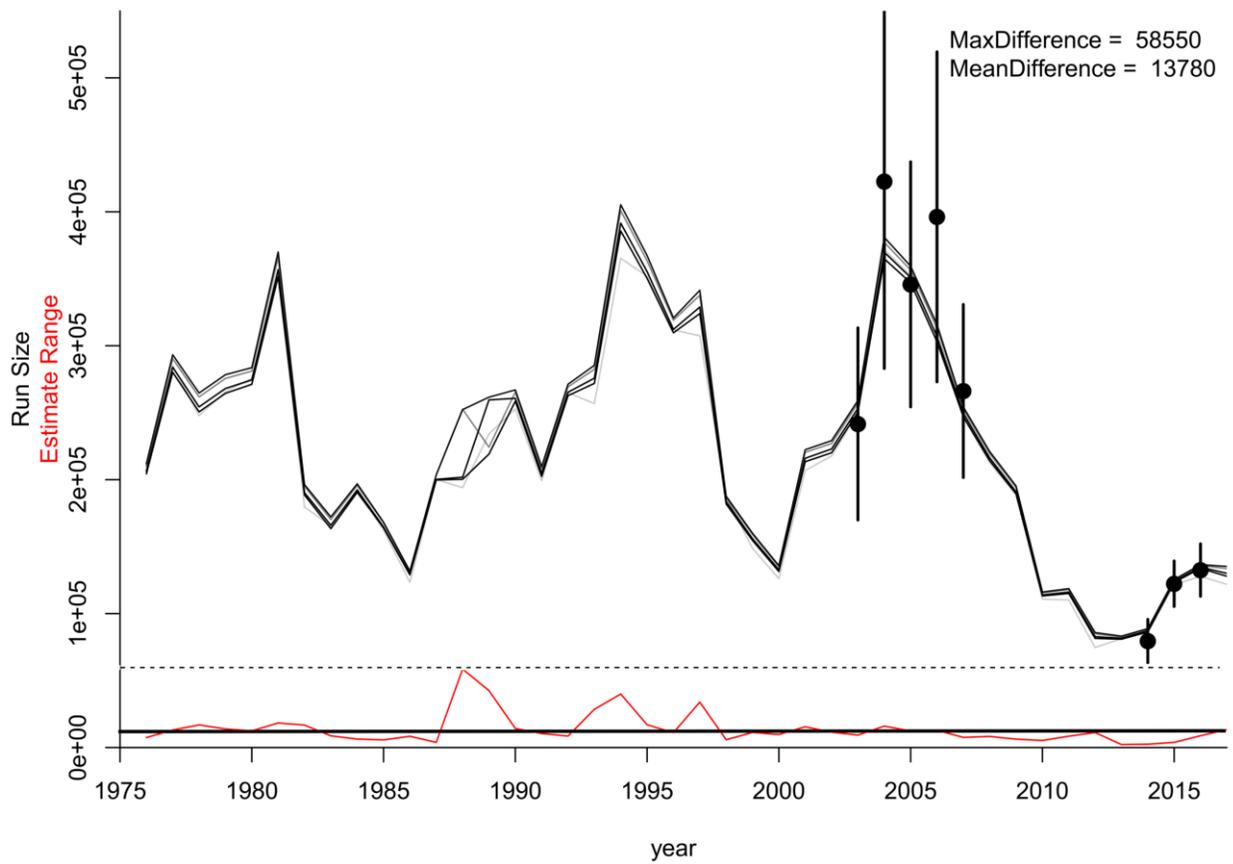


Figure 12. As in Fig. 7, but with over-dispersion parameters pooled by escapement index type (weir/aerial).

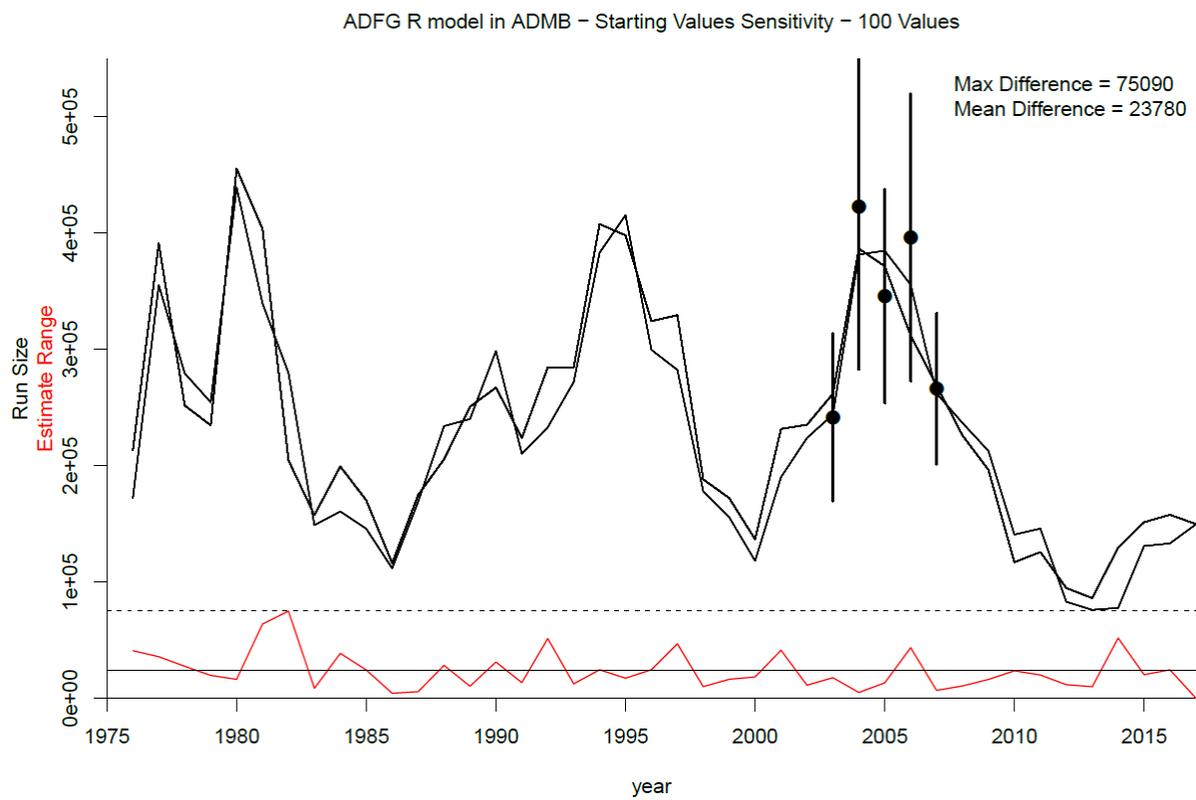


Figure 13. As in Fig. 5, but removing the harvest component from the likelihood function.

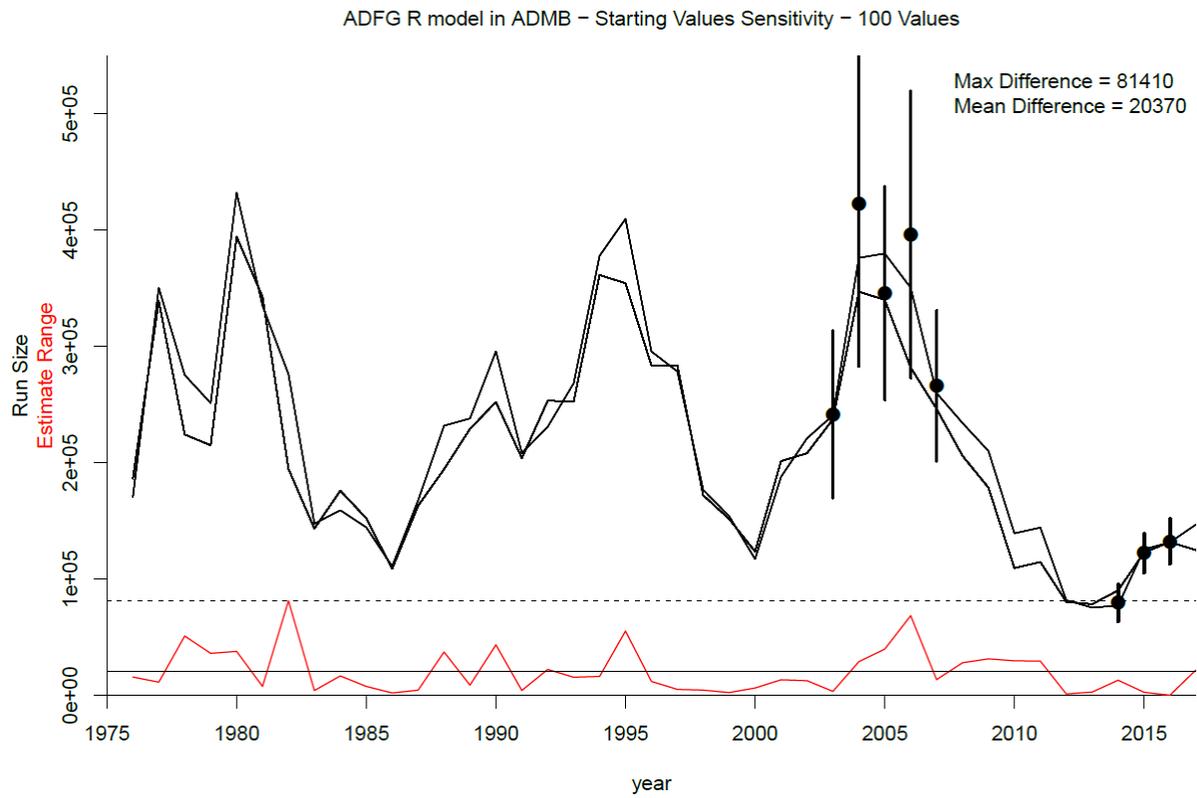


Figure 14. As in Fig. 7, but removing the harvest component from the likelihood function.

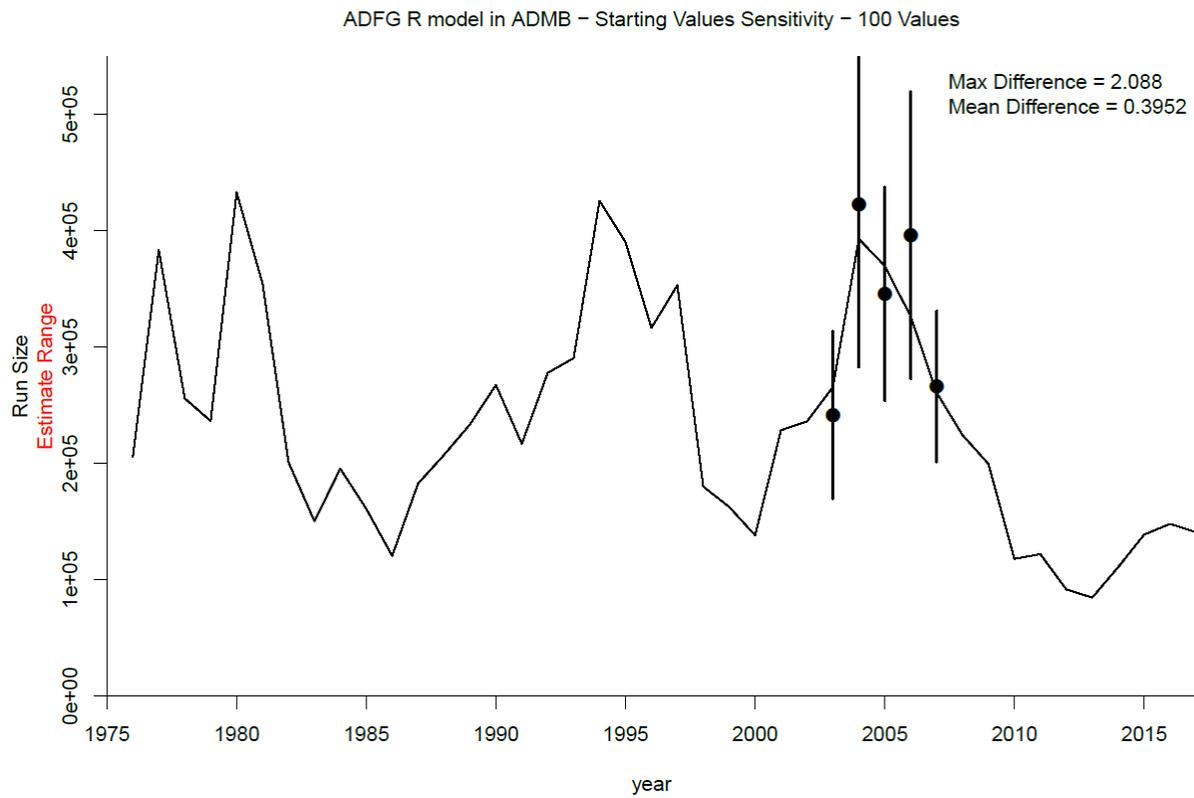


Figure 15. As in Fig. 10, but removing the harvest component from the likelihood function.

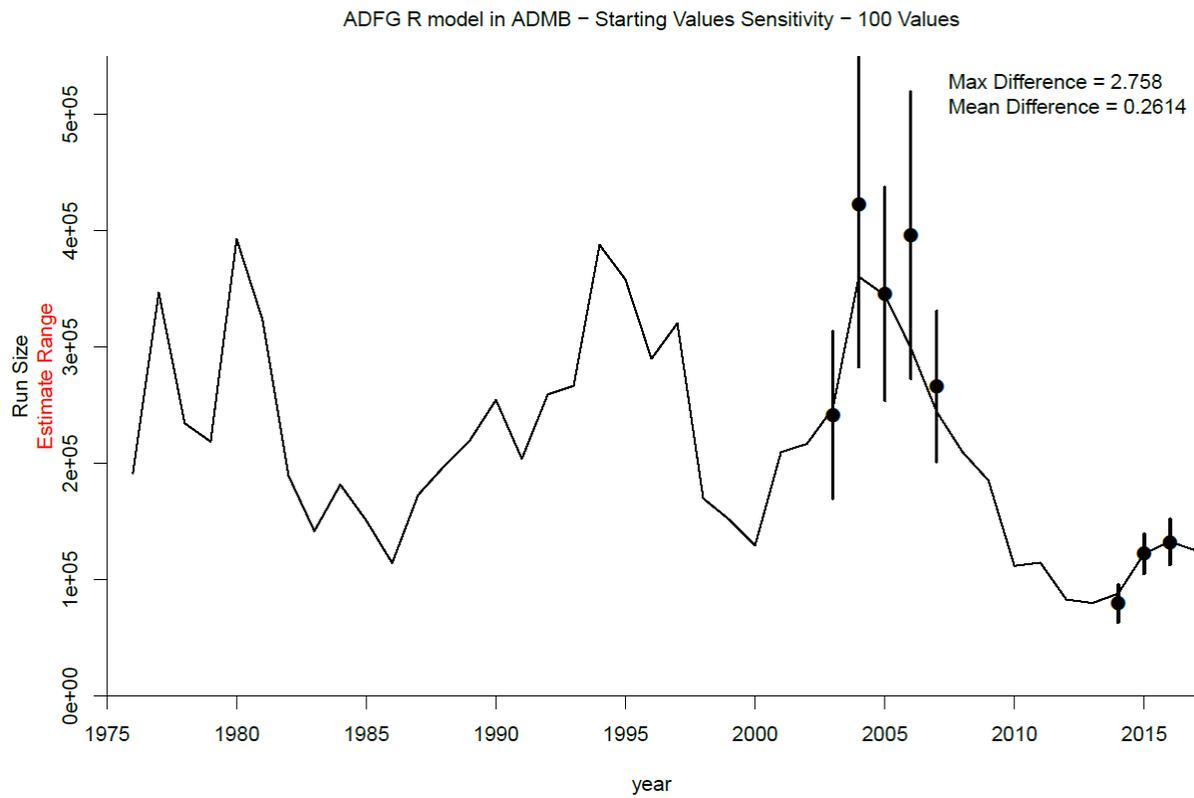


Figure 16. As in Fig. 12, but removing the harvest component from the likelihood function.

R model in ADMB, Individual Dispersion, With new MR, Perfect Synchrony, .7 auto, no ObsErr, no harv

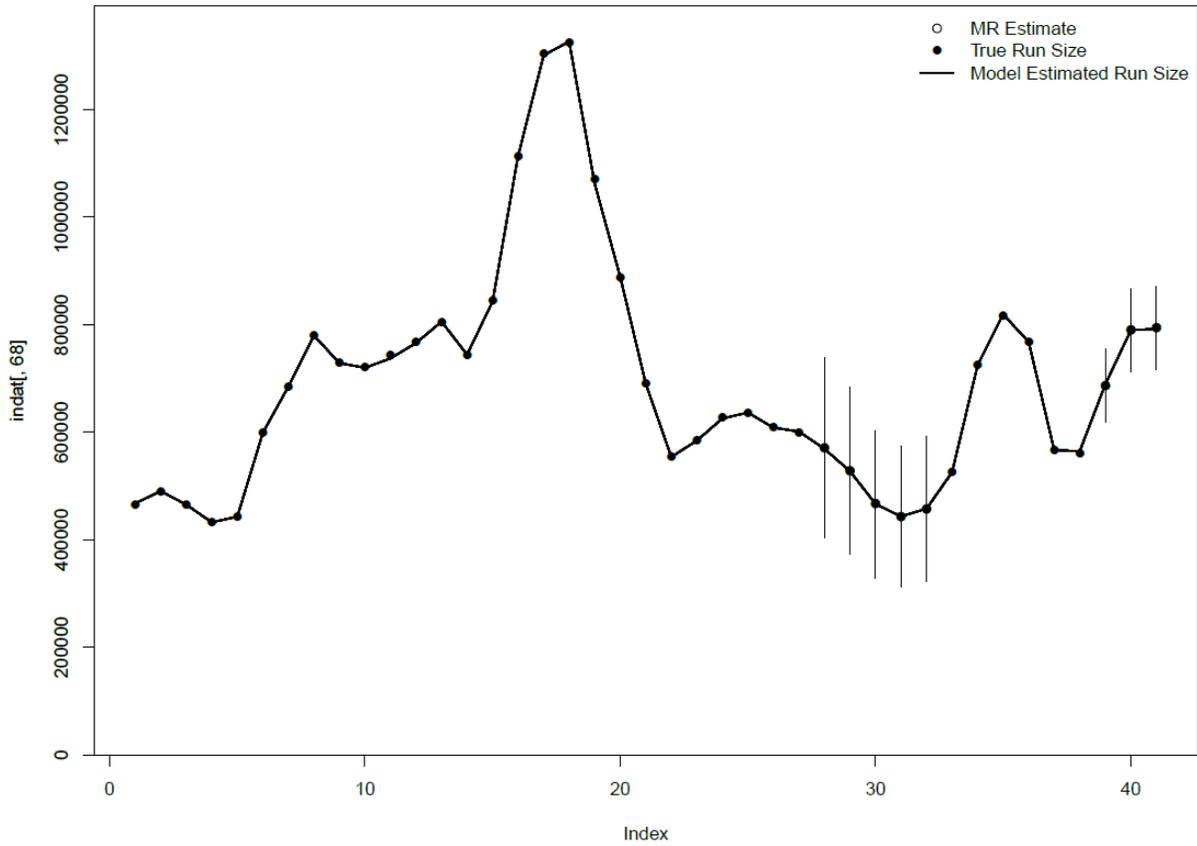


Figure 17. Original R model in ADMB fit to simulated data without the harvest component of the likelihood. This scenario had new mark-recapture estimates incorporated, perfect synchrony among sub-stocks, and no observation error. Black points represent true run size values in the simulation model. Points with error bars represent mark-recapture estimates. The thick black line is composed of 100 individual model fits across a range of starting values.

R model in ADMB, Individual Dispersion, With new MR, Perfect Synchrony, .7 auto, with ObsErr, no harv



Figure 18. As in Fig. 17, but with observation error in the aerial survey and weir indices.

R model in ADMB, Individual Dispersion, With new MR, 0 Synchrony, .7 auto, no ObsErr, no harv

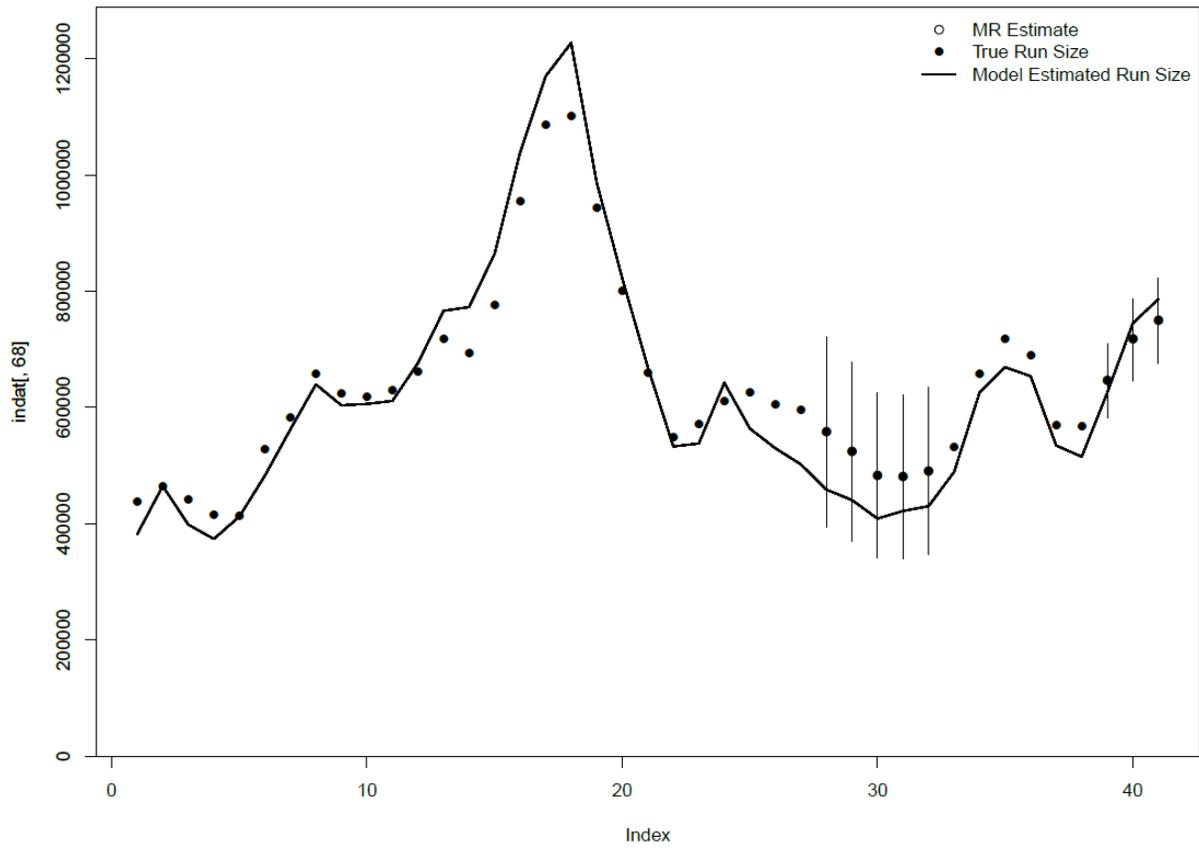


Figure 19. As in Fig. 17, but with asynchronous sub-population dynamics.

R model in ADMB, Individual Dispersion, With new MR, 0 Synchrony, .7 auto, with ObsErr, no harv

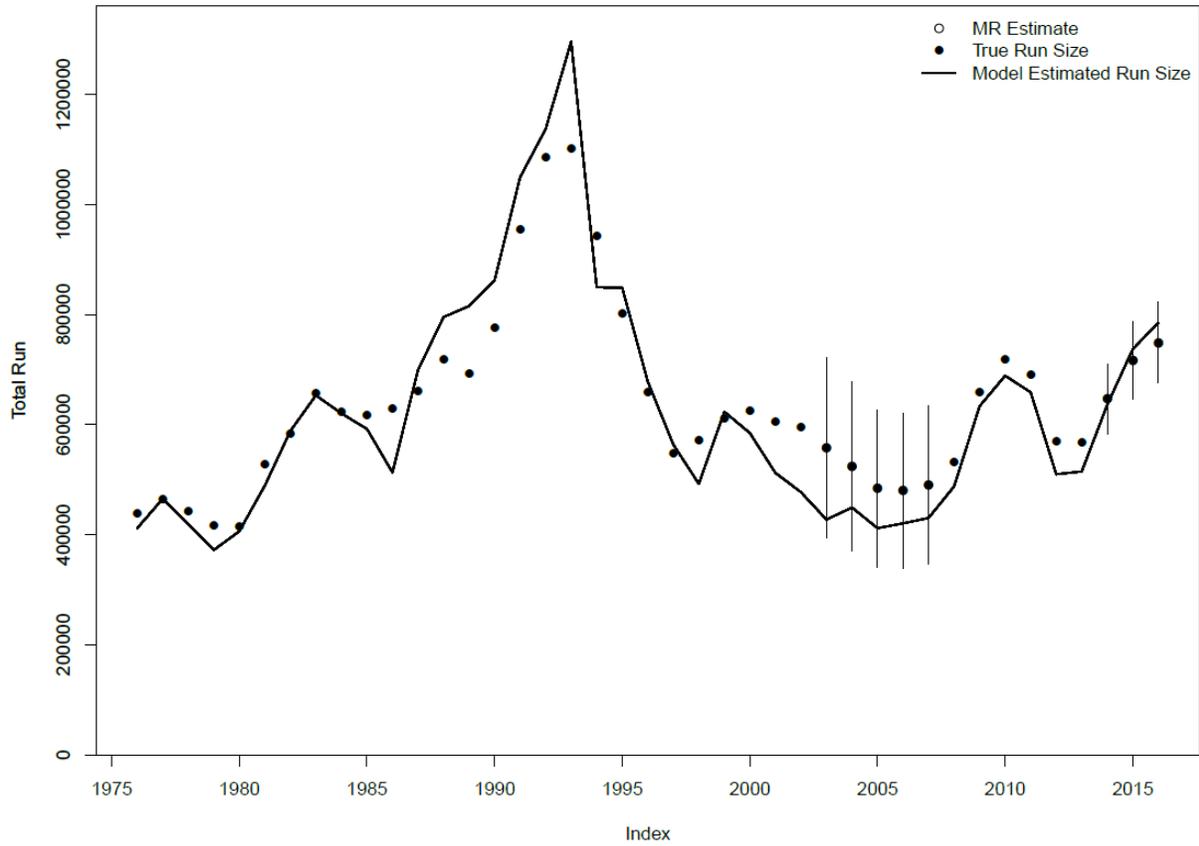


Figure 20. As in Fig. 18, but with asynchronous sub-population dynamics.

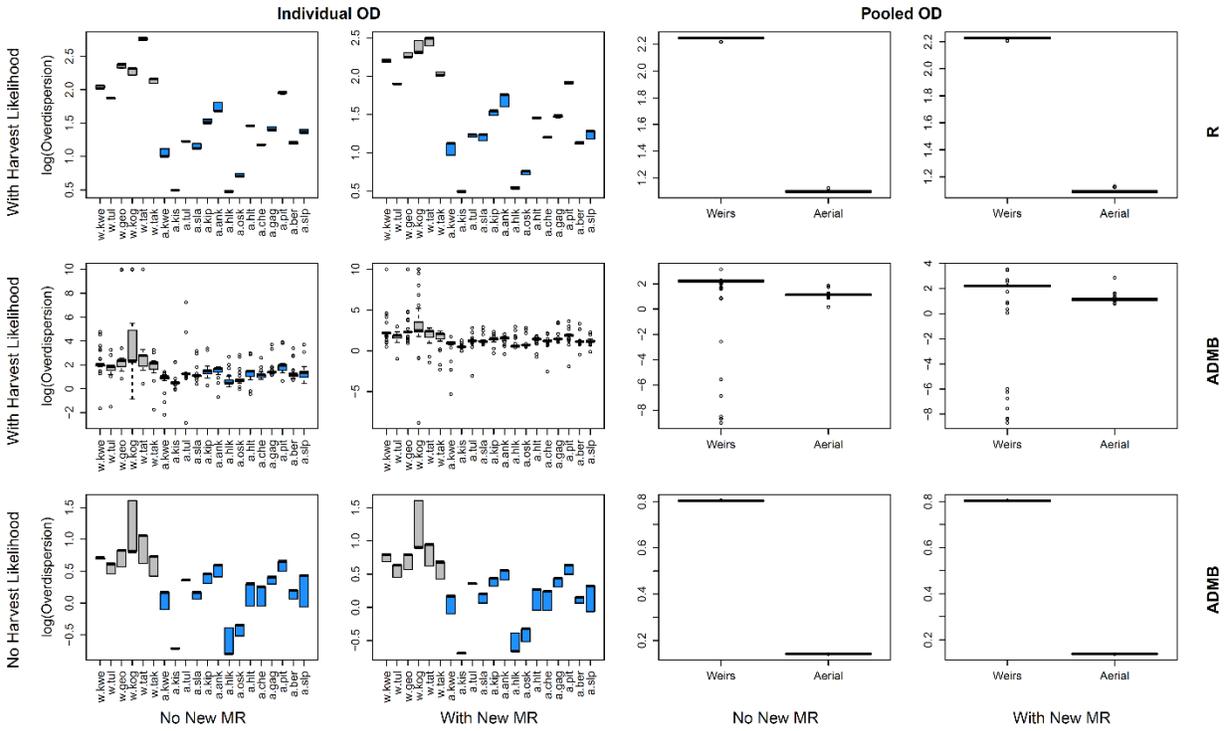


Figure 22. Model estimates of escapement-index over-dispersion parameters across different model structures and levels of data availability. Boxes indicate over-dispersion parameters that were estimated for weir (grey) and aerial survey (blue) indices. The top row shows results for models fit in R, while the bottom two rows show results of models fit in ADMB. The top two rows show the results of models with the harvest component of the likelihood included, while the bottom row shows models fit without the harvest component of the likelihood. The left two columns of panels are model fits with individual over-dispersion parameters for each escapement index (with those beginning with a “w.” being weir projects, and those beginning with an “a.” being aerial surveys), while the right two columns of panels show model fits with over-dispersion parameters pooled by escapement index type. The first and third columns show model fits when no new mark-recapture data are available, while the second and fourth columns show model fits with new mark recapture data available. Boxes represent the interquartile range of parameter estimates, thick line inside of box indicates the median estimate, whiskers extend to the furthest estimate within 1.5x the IQR of the median, and points represent all estimates that are more than 1.5x the IQR less or greater than the median estimate.

Individual OD Parameters

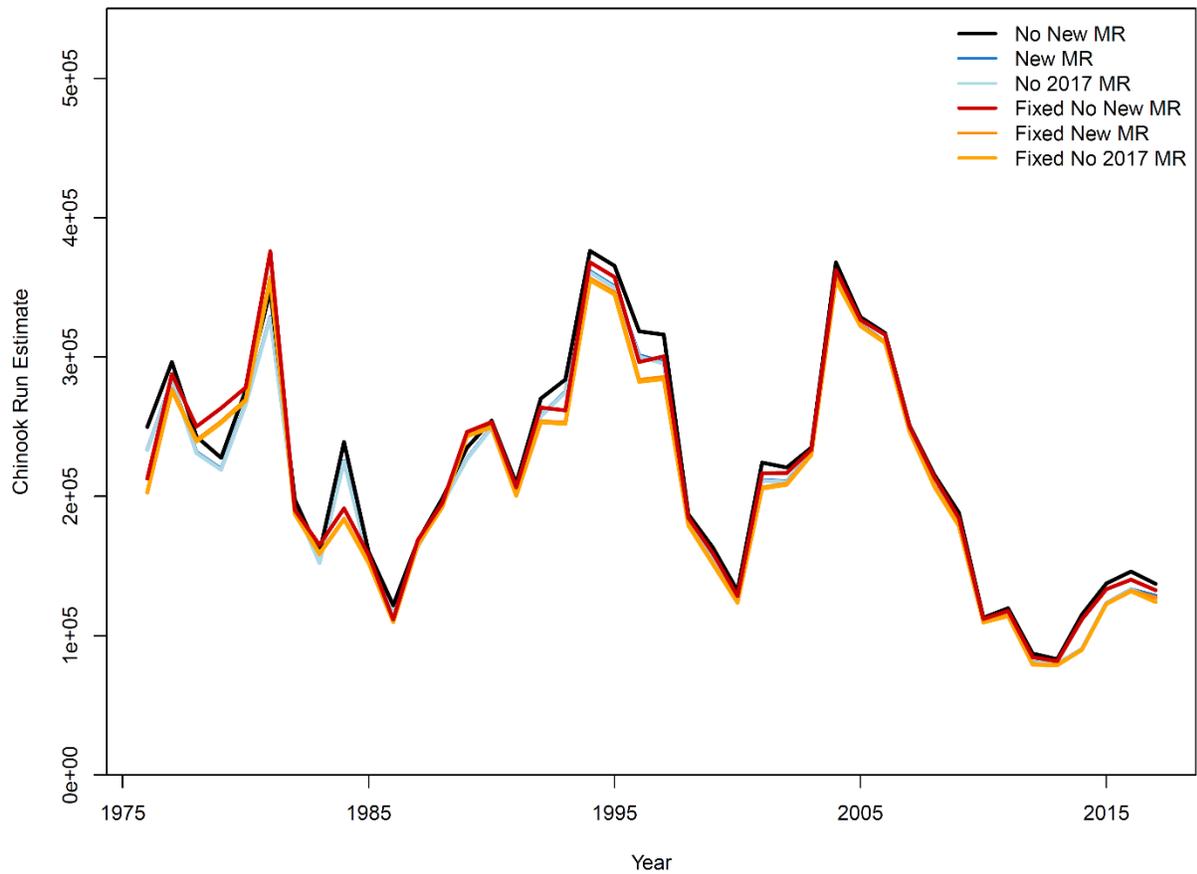


Figure 23. Effects of pooling catch and effort data across the final three weeks of season on run-reconstruction estimates. Black line is the base model, while blue lines are the base model with new mark-recapture estimates included. The orange lines are fits when catch and effort are not pooled across the final three weeks of the season. All of these models have individual over-dispersion parameters for each escapement index.

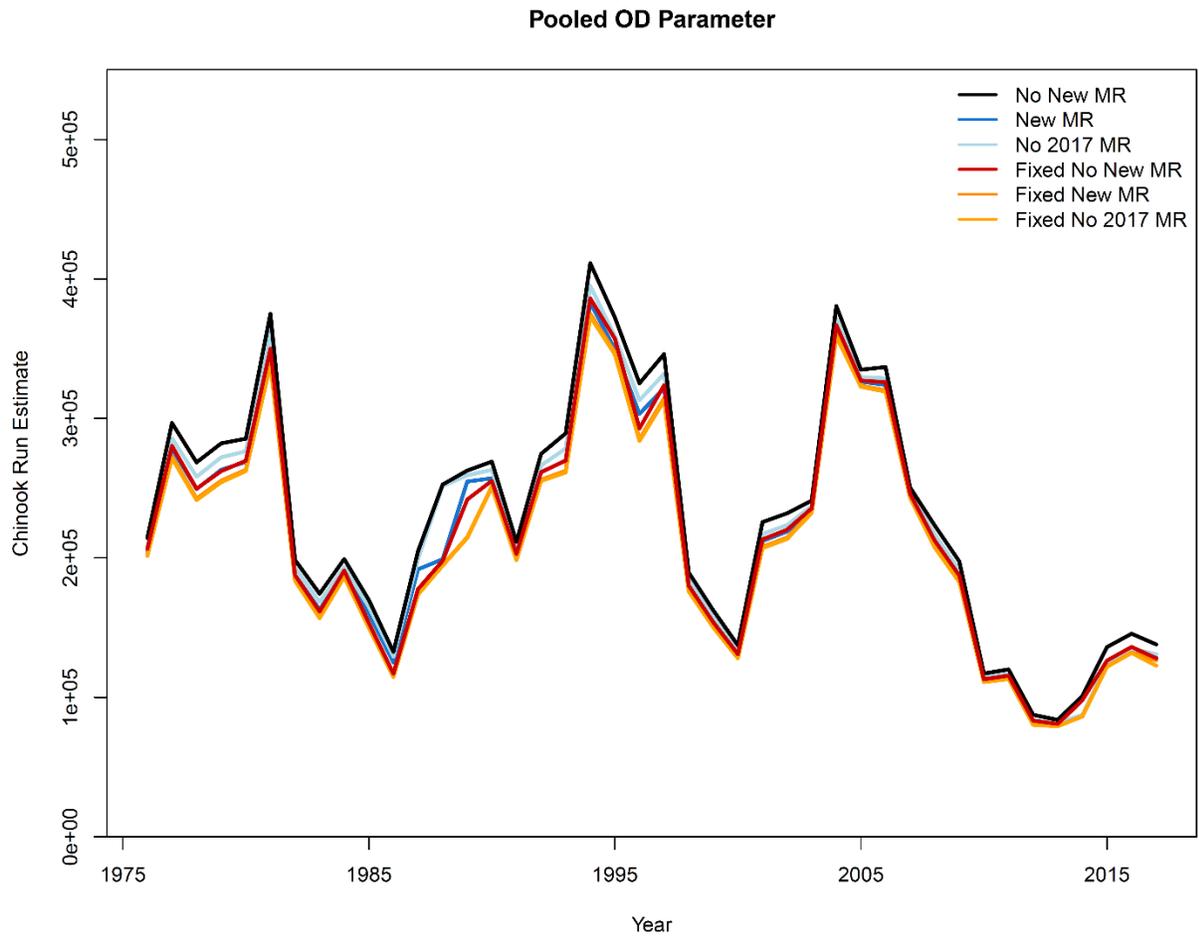


Figure 24. As in Fig. 23, but all of these models have over-dispersion parameters pooled by escapement index type (weir/aerial survey).

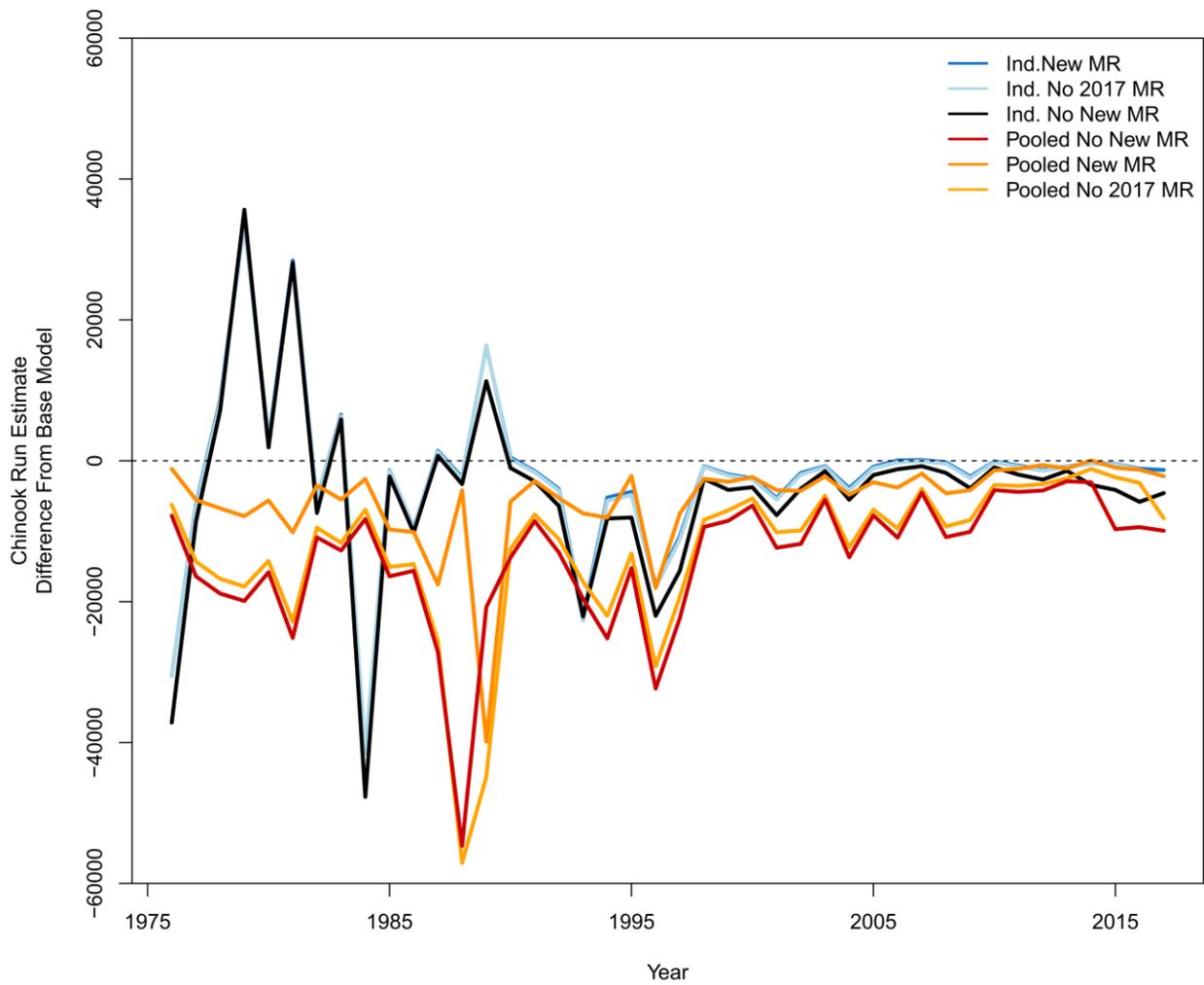


Figure 25. Change in run-reconstruction estimates from no longer pooling catch and effort data across the final three weeks of the season. Each model structure in this figure is compared to the same model structure, but with pooled catch and effort.

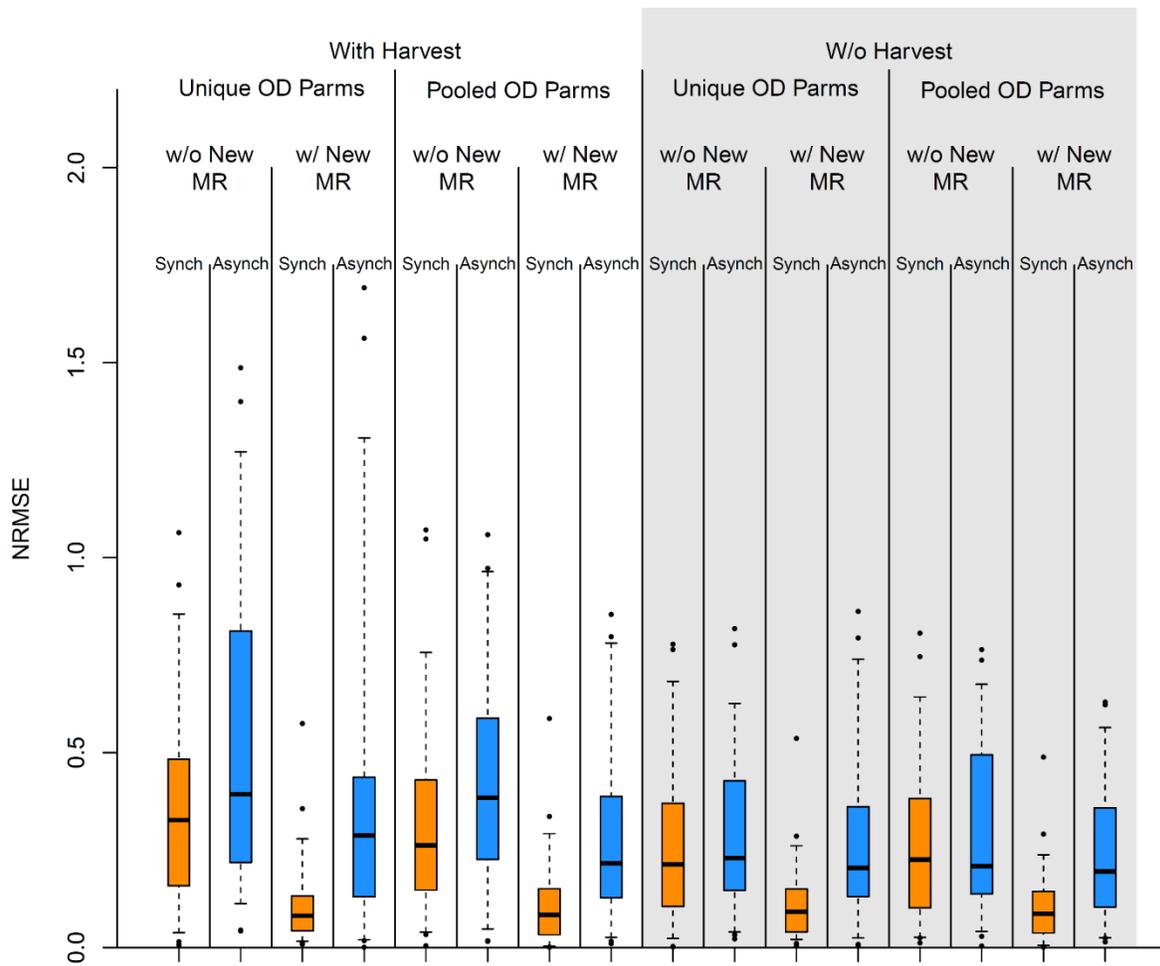


Figure 26. Boxplots of normalized RMSE for Kuskokwim Run-reconstruction model fits to simulated data under different biological scenarios and model structures. The boxes show the interquartile range (IQR) and the whiskers extend to the most extreme value within the 95 % simulation interval. The colors represent scenarios with synchronous (orange) and asynchronous (blue) sub-stock dynamics. Column labels describe whether or not new mark-recapture estimates are available when fitting the model, whether or not escapement over-dispersion parameters are pooled in the run-reconstruction model, and whether or not the harvest likelihood is included in the run-reconstruction model (the latter assumption also indicated by the grey background).

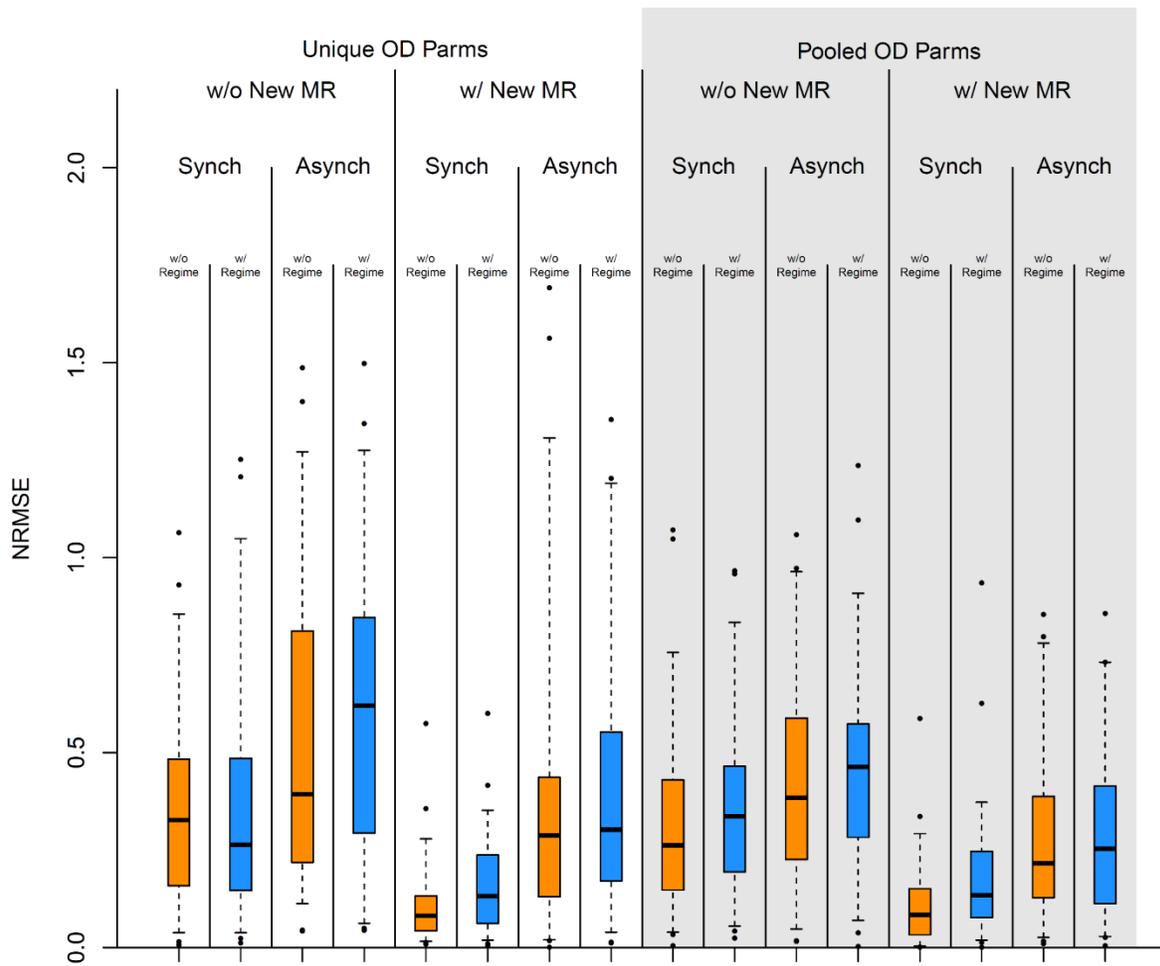


Figure 27. Boxplots of normalized RMSE for Kuskokwim Run-reconstruction model fits to simulated data under different biological scenarios and model structures. The colors represent scenarios with (blue) and without (orange) regime shifts in population productivity. Column labels describe whether or not sub-populations demonstrate perfect synchrony or asynchronous dynamics, whether or not new mark-recapture estimates are available when fitting the model, and whether or not escapement over-dispersion parameters are pooled in the run-reconstruction model (also indicated by the grey background).

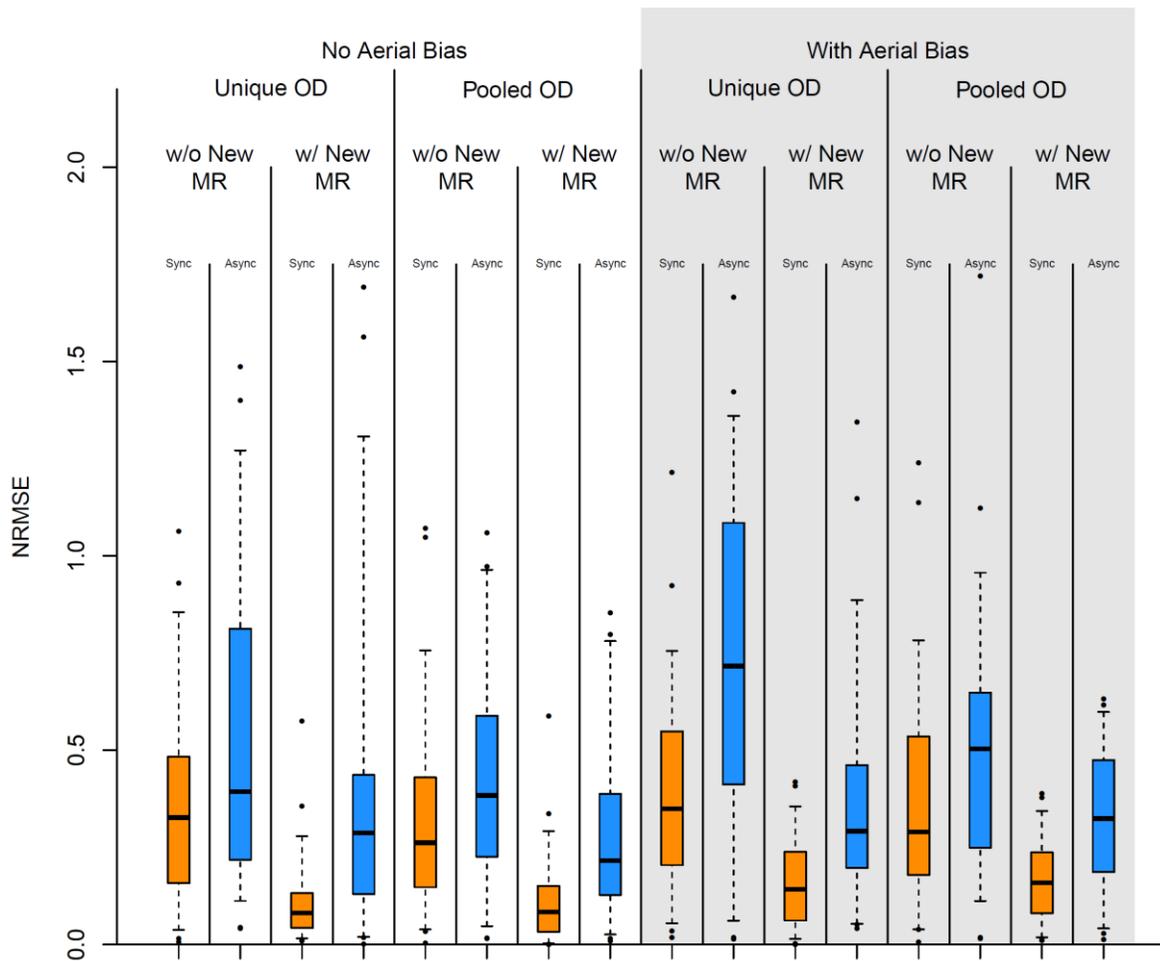


Figure 28. Boxplots of normalized RMSE for Kuskokwim Run-reconstruction model fits to simulated data under different biological scenarios and model structures. The colors represent scenarios with synchronous (orange) and asynchronous (blue) sub-stock dynamics. Column labels describe whether or not new mark-recapture estimates are available when fitting the model, whether or not escapement over-dispersion parameters are pooled in the run-reconstruction model, and whether or not aerial escapement indices are systematically biased because of the assumption of linearity (also indicated by the grey background).

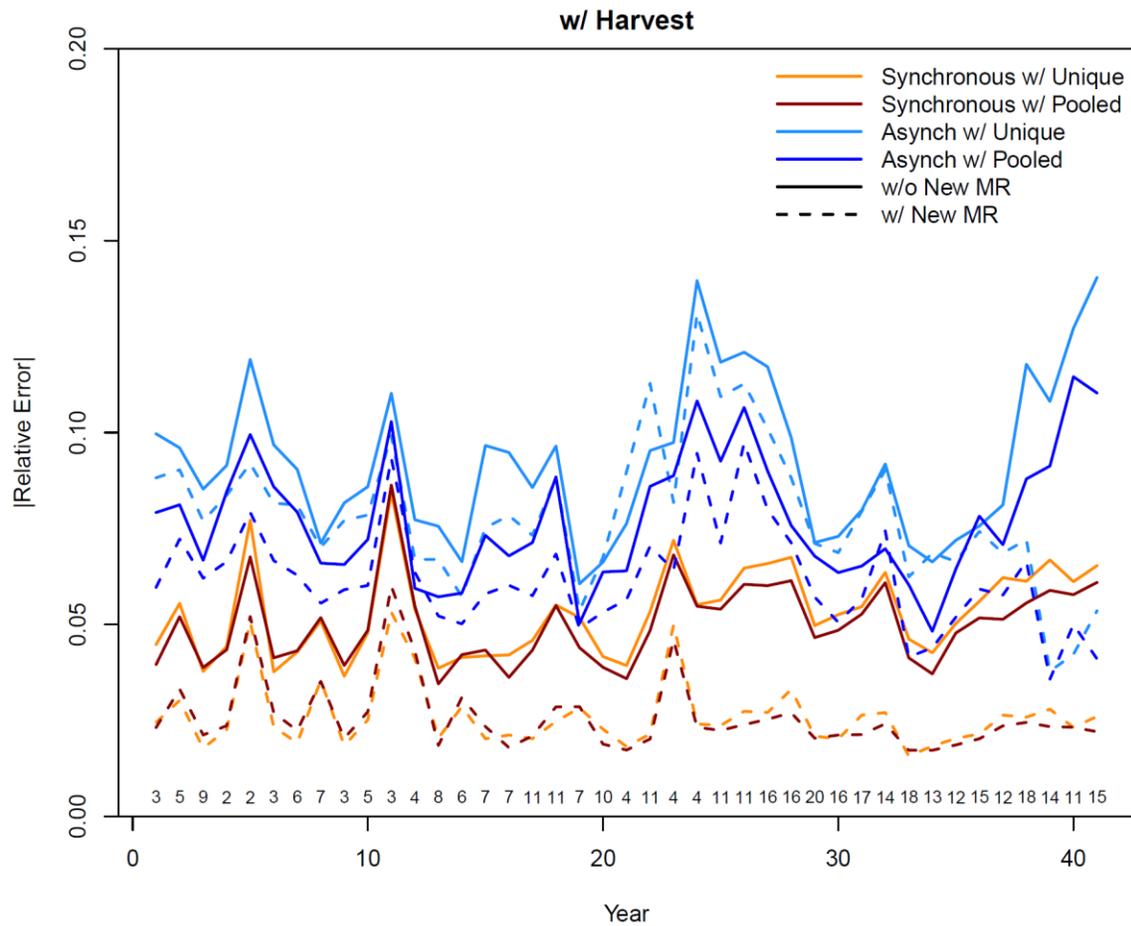


Figure 29. Absolute values of relative error through time in run-reconstruction estimates for simulated time-series where the harvest component of the likelihood was included. Solid lines represent those in which the new mark-recapture estimates are not available for the run-reconstruction model. Dashed lines represent scenarios in which the new mark-recapture estimates are available for the run-reconstruction model. Lines in orange shades represent scenarios where the underlying sub-populations are perfectly synchronous, while blue shaded lines represent those with asynchronous sub-populations. Darker shades of each color represent scenarios with pooled over-dispersion parameters in the run-reconstruction model, while lighter shades represent scenarios with unique over-dispersion parameters for each escapement index. Numbers above x-axis indicate the number of escapement indices available each year.

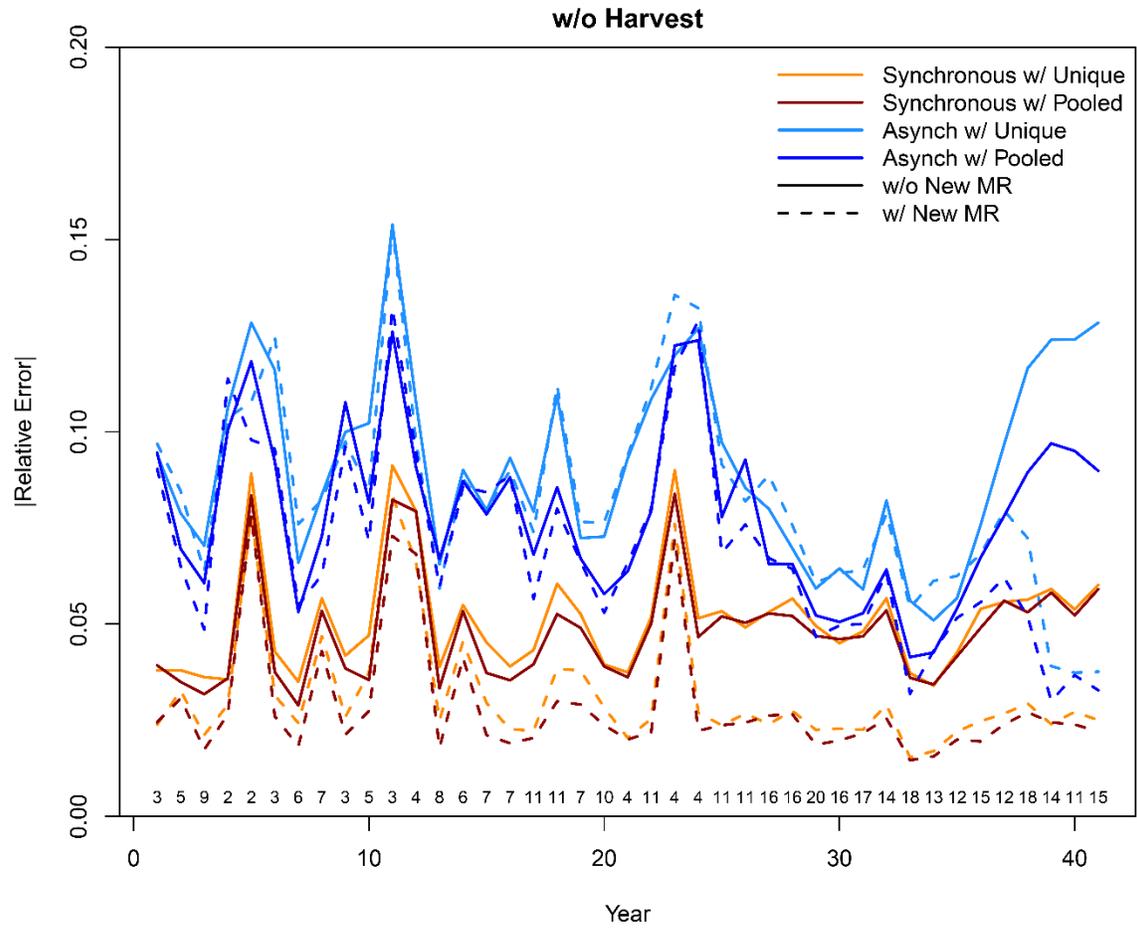


Figure 30. As for Fig. 29, except that the harvest component of the likelihood is excluded.

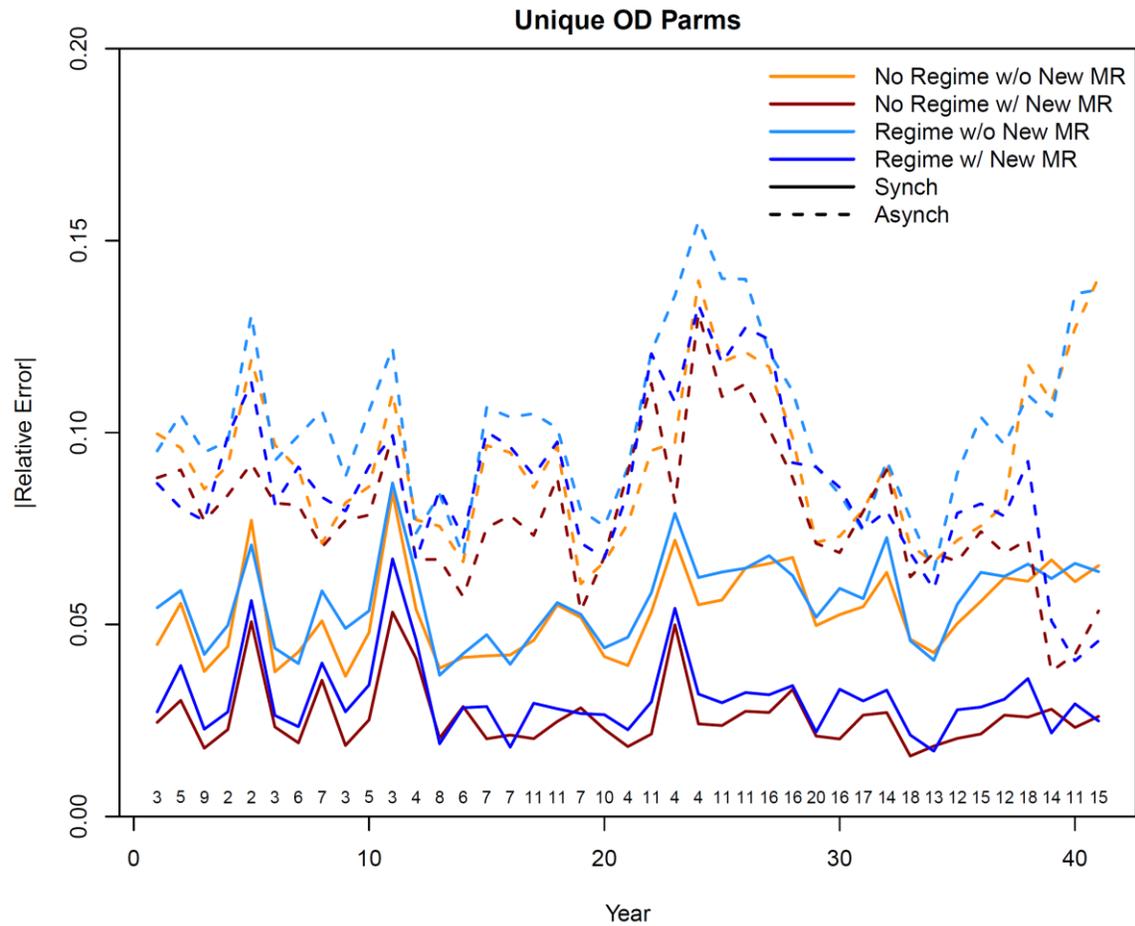


Figure 31. Absolute values of bias through time in run-reconstruction estimates for simulated time-series where the estimation model had unique over-dispersion parameters for each escapement index. Solid lines represent those in which the sub-population dynamics are perfectly synchronous. Dashed lines represent scenarios in which the sub-population dynamics are asynchronous. Lines in orange shades represent scenarios where no regime shifts occur, while blue shaded lines represent those with regime shifts. Darker shades of each color represent scenarios with new mark-recapture estimates available to the run reconstruction model, while lighter shades represent scenarios without new mark-recapture estimates being available. Numbers above x-axis indicate the number of escapement indices available each year.

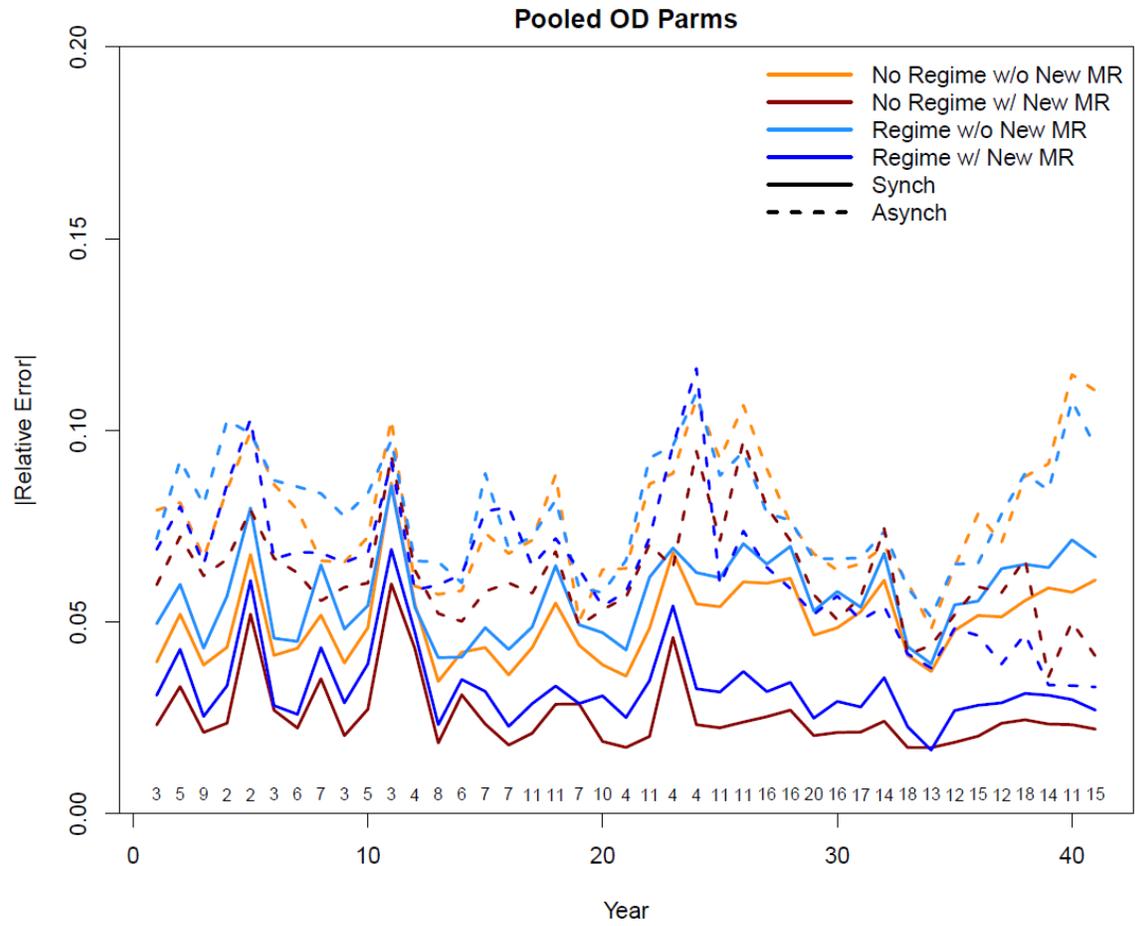


Figure 32. As in Fig. 31, but with over-dispersion parameters pooled by escapement index type (weir/aerial survey).

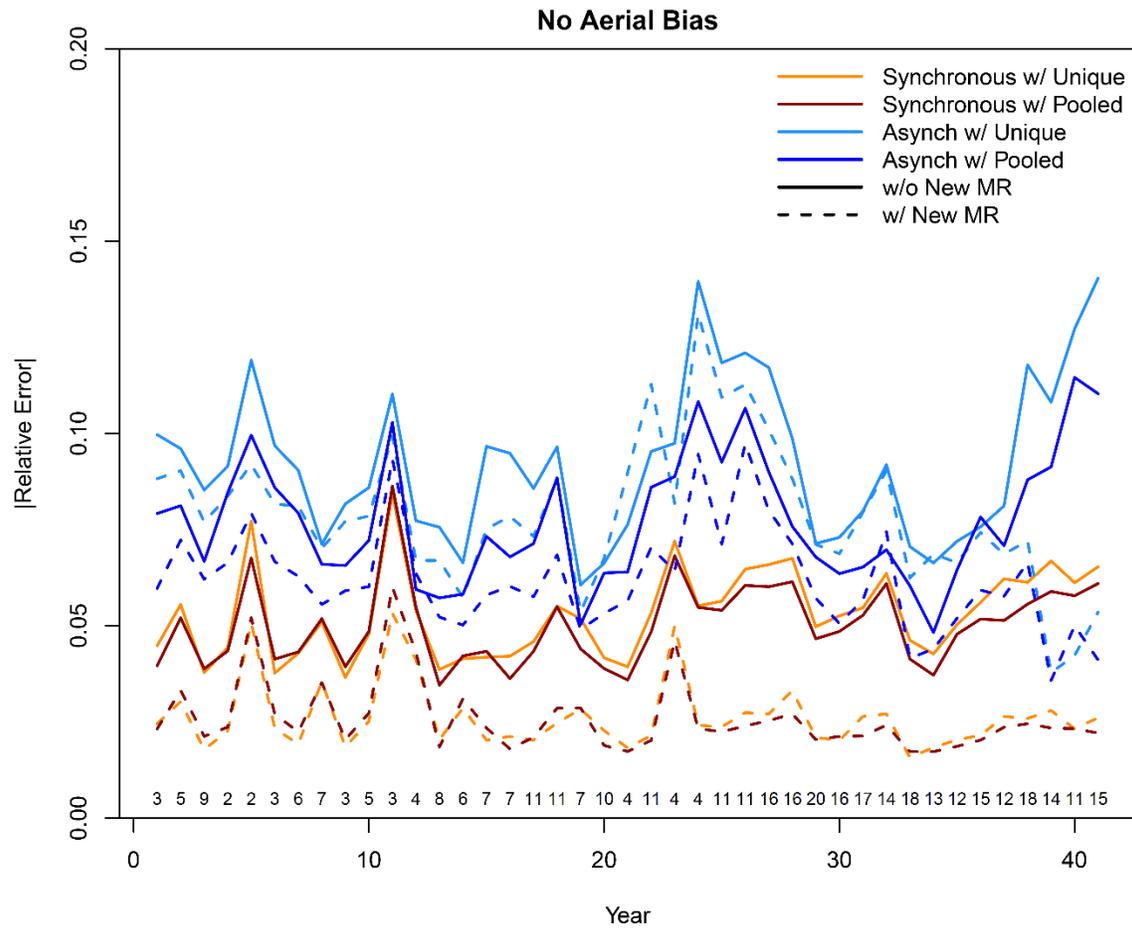


Figure 33. Absolute values of bias through time in run-reconstruction estimates for simulated time-series where there is no systematic bias in aerial survey indices. Solid lines represent those in which the new mark-recapture estimates are not available for the run-reconstruction model. Dashed lines represent scenarios in which the new mark-recapture estimates are available for the run-reconstruction model. Lines in orange shades represent scenarios where the underlying sub-populations demonstrate perfectly synchronous dynamics, while blue shaded lines represent those with asynchronous sub-populations. Darker shades of each color represent scenarios with pooled over-dispersion parameters in the run-reconstruction model, while lighter shades represent scenarios with unique over-dispersion parameters for each escapement index. Numbers above x-axis indicate the number of escapement indices available each year.

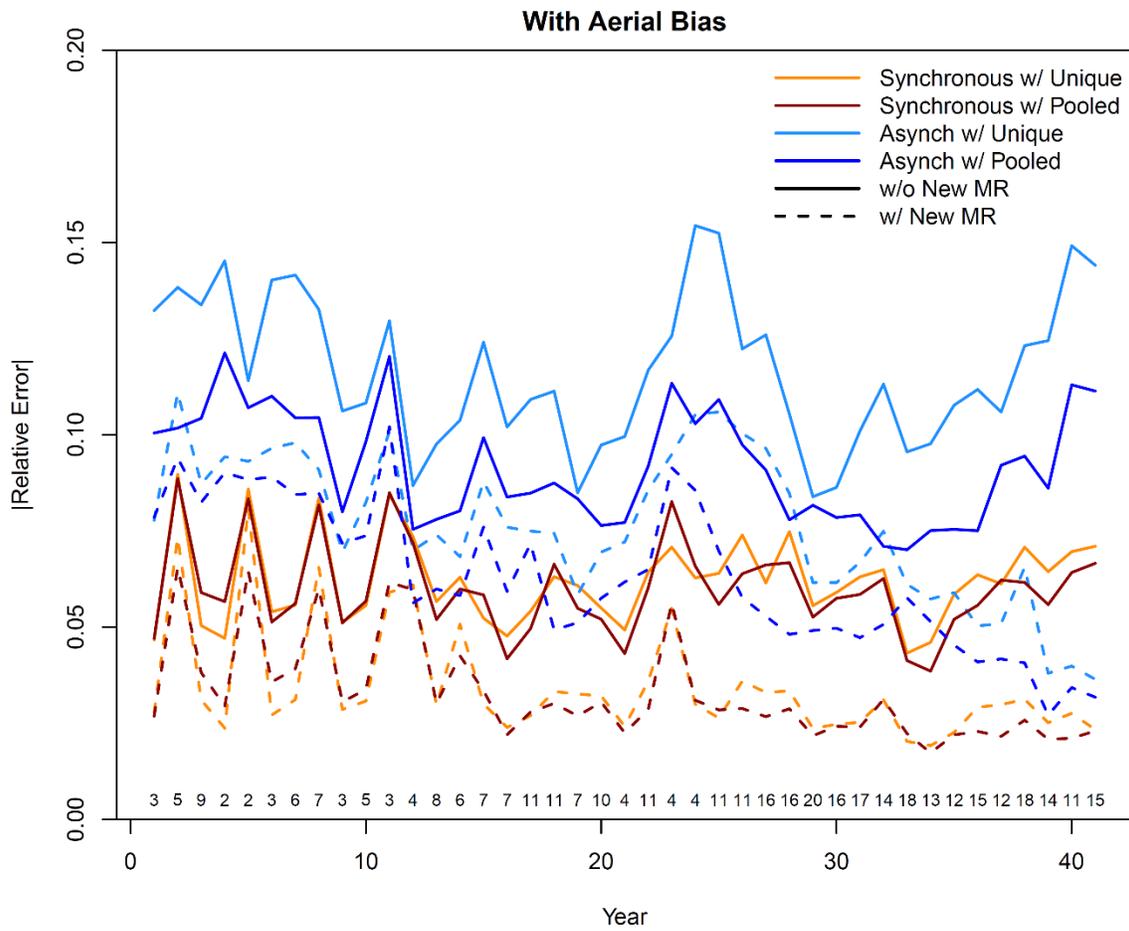


Figure 34. As in Fig. 33, but with non-linear bias in aerial survey indices.

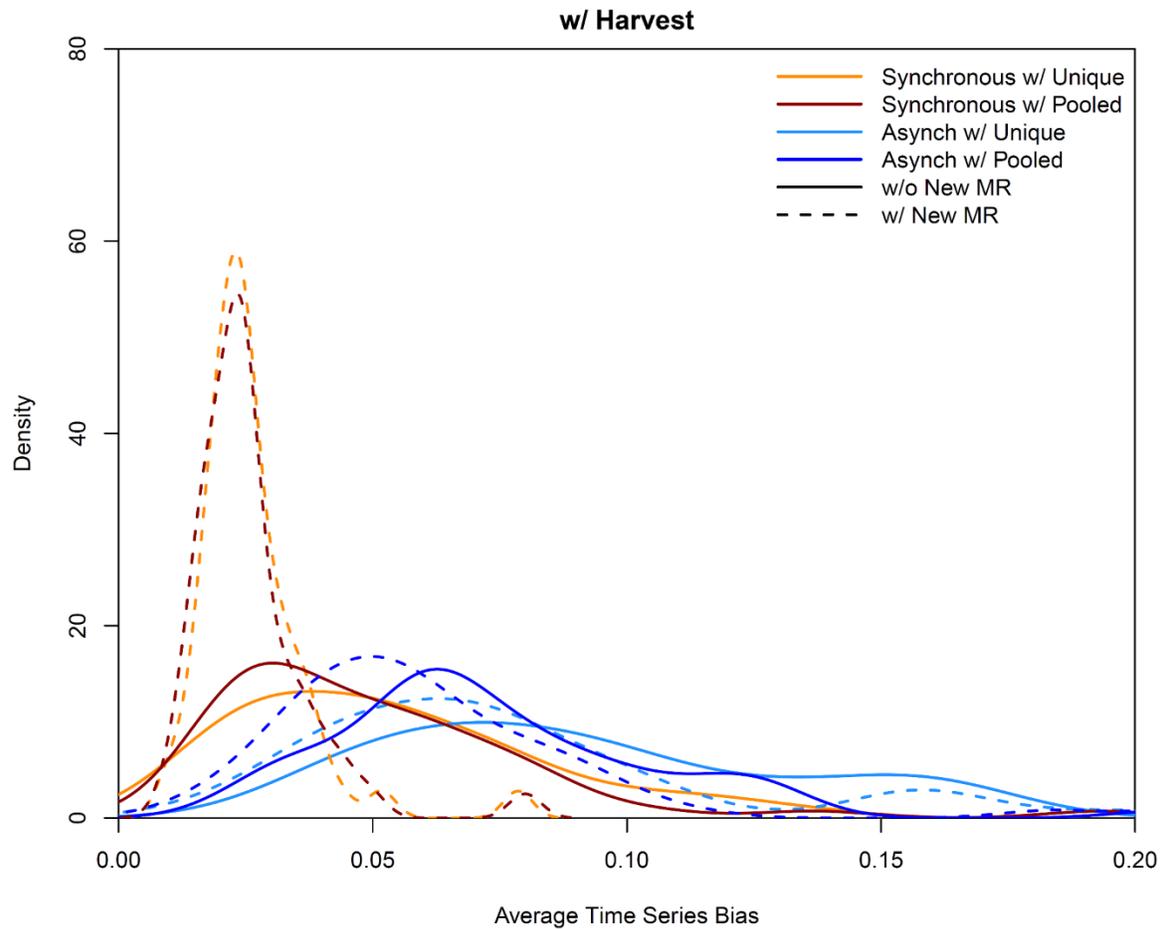


Figure 35. Density plots of the absolute value of the average time-series bias across all iterations for different scenarios for both the operating model and estimation model, where the harvest component of the likelihood is included in the estimation model. Higher values indicate more systematic bias across an entire time-series. Solid lines represent those in which the new mark-recapture estimates are not available for the run-reconstruction model. Dashed lines represent scenarios in which the new mark-recapture estimates are available for the run-reconstruction model. Lines in orange shades represent scenarios where the underlying sub-populations demonstrate perfectly synchronous dynamics, while blue shaded lines represent those with asynchronous sub-populations. Darker shades of each color represent scenarios with pooled over-dispersion parameters in the run-reconstruction model, while lighter shades represent scenarios with unique over-dispersion parameters for each escapement index.

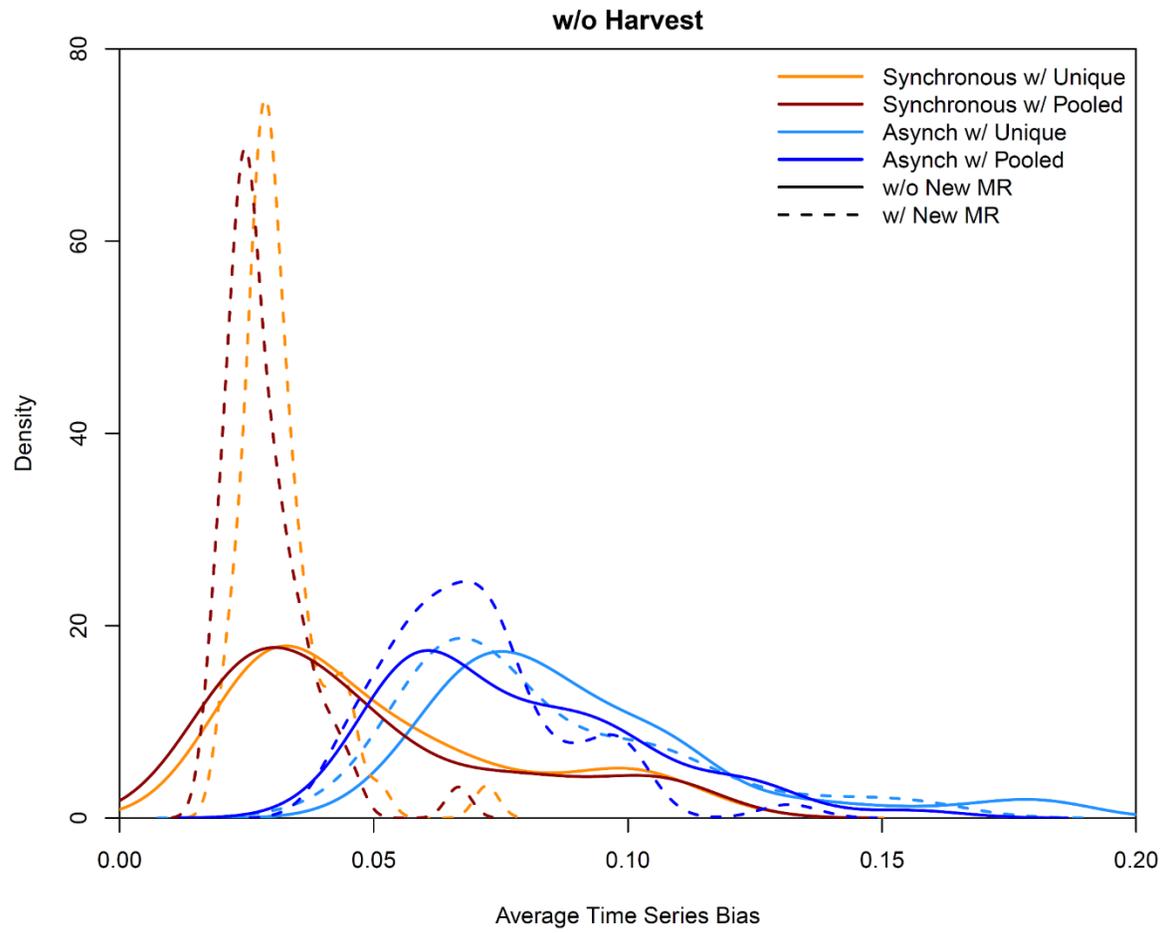


Figure 36. As in Fig. 35, but without the harvest component of the likelihood.

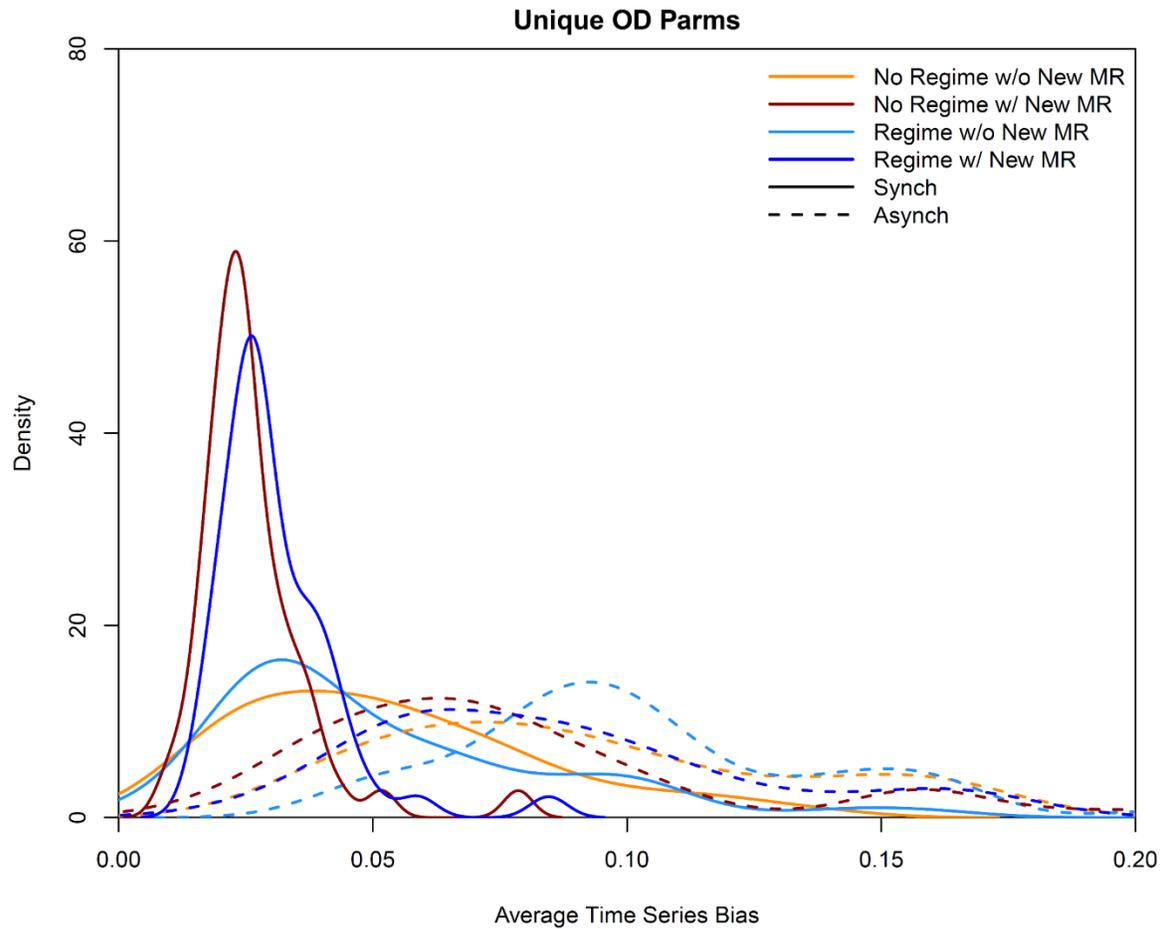


Figure 37. Density plots of the absolute value of the average time-series bias across all iterations for different scenarios for both the operating model and estimation model, where the over-dispersion parameters in the estimation model were unique for each escapement index. Higher values indicate more systematic bias across an entire time-series. Solid lines represent those in which the sub-population dynamics are perfectly synchronous. Dashed lines represent scenarios in which the sub-population dynamics are asynchronous. Lines in orange shades represent scenarios where no regime shifts occur, while blue shaded lines represent those with regime shifts. Darker shades of each color represent scenarios with new mark-recapture estimates available to the run-reconstruction model, while lighter shades represent scenarios without new mark-recapture estimates being available.

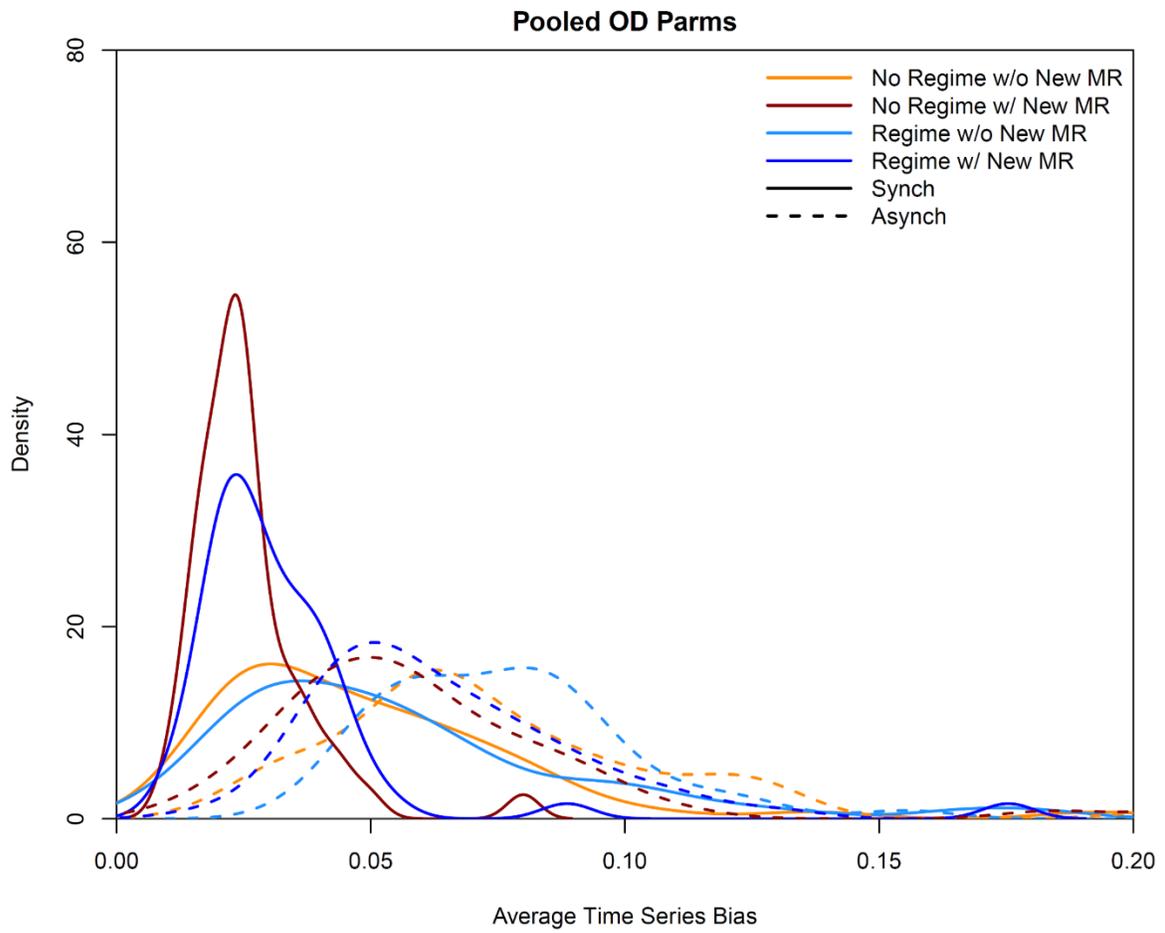


Figure 38. As in Fig. 37, but with over-dispersion parameters pooled by escapement index type (weir/aerial survey).

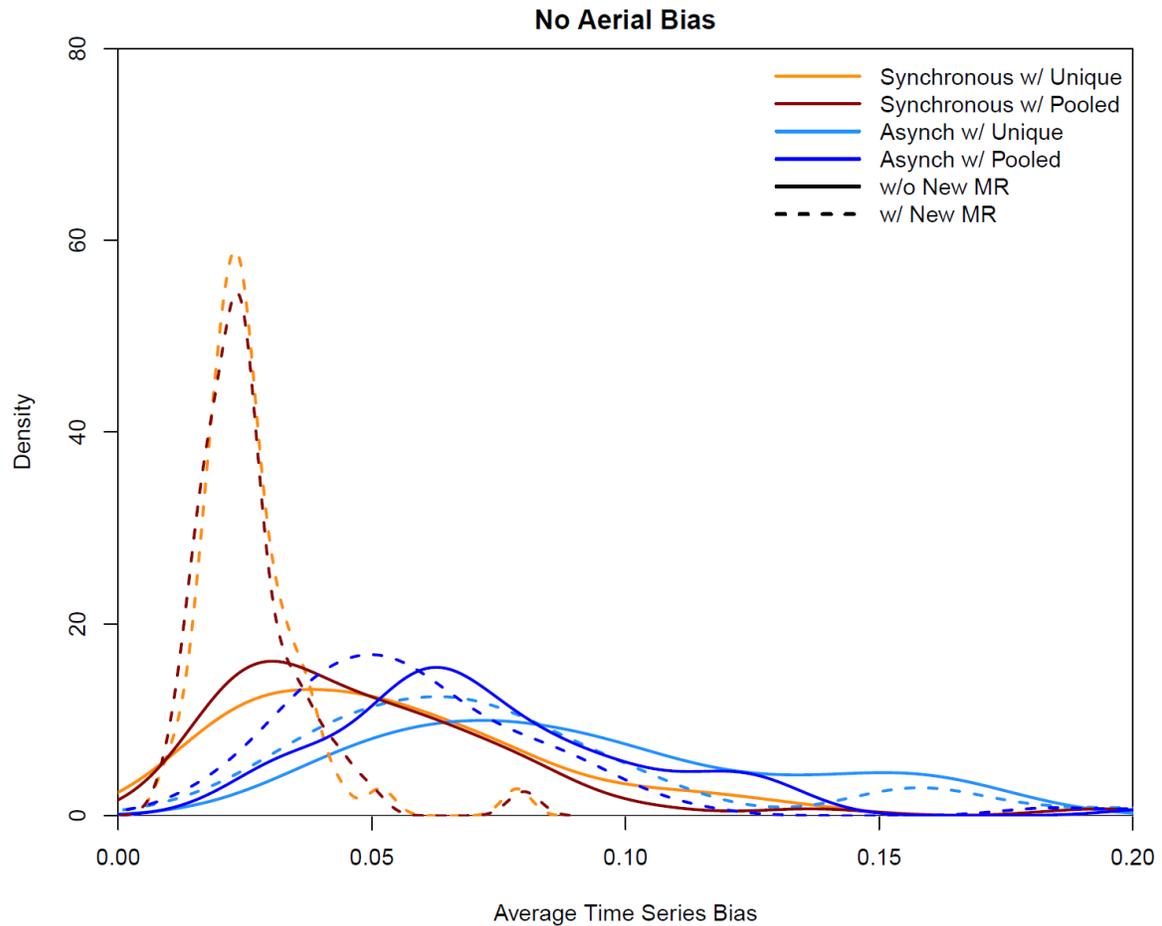


Figure 39. Density plots of the absolute value of the average time-series bias across all iterations for different scenarios for both the operating model and estimation model, where there is no systematic bias in aerial survey indices. Higher values indicate more systematic bias across an entire time-series. Solid lines represent those in which the new mark-recapture estimates are not available for the run-reconstruction model. Dashed lines represent scenarios in which the new mark-recapture estimates are available for the run-reconstruction model. Lines in orange shades represent scenarios where the underlying sub-populations demonstrate perfectly synchronous dynamics, while blue shaded lines represent those with asynchronous sub-populations. Darker shades of each color represent scenarios with pooled over-dispersion parameters in the run-reconstruction model, while lighter shades represent scenarios with unique over-dispersion parameters for each escapement index.

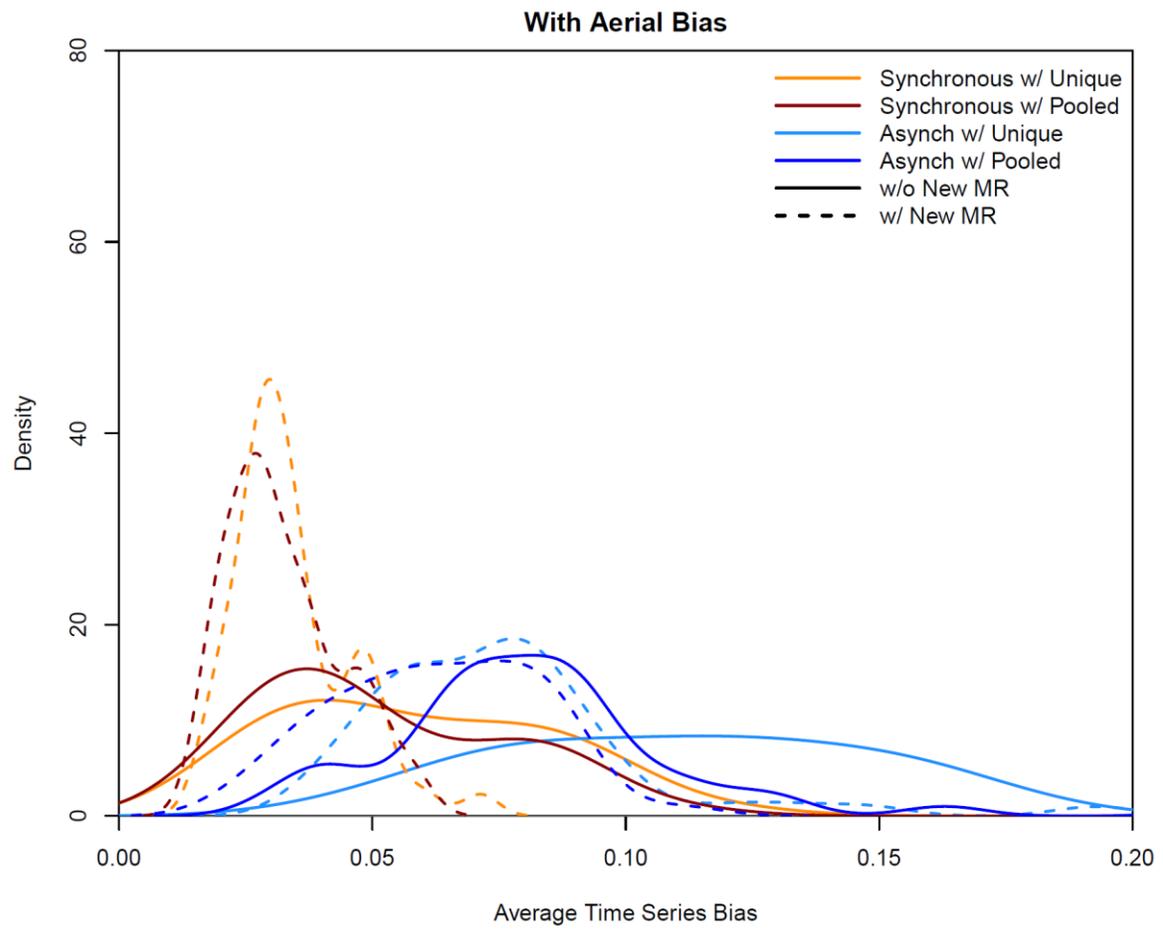


Figure 40. As in Fig. 39, but with non-linear bias in aerial survey indices.

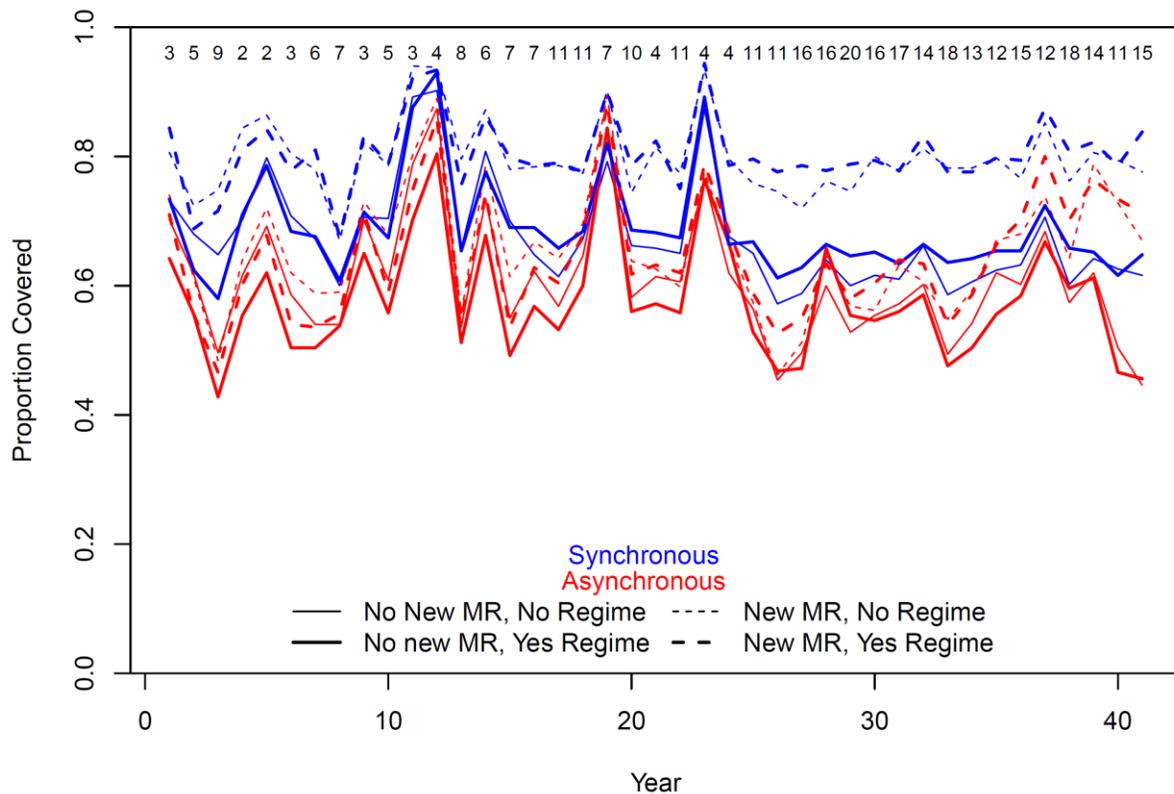


Figure 41. The proportion of simulations in which the true run-size value was within the 95% confidence interval of the model estimated run-size across time under different scenarios. Blue (red) lines indicate scenarios with synchronous (asynchronous) sub-population dynamics. Solid (dashed) lines indicate scenarios without (with) the new mark-recapture estimates available. Thin (thick) lines indicate scenarios without (with) regime shifts impacting the underlying population dynamics. The numbers at the top indicate the number of escapement indices available to the model in that year.

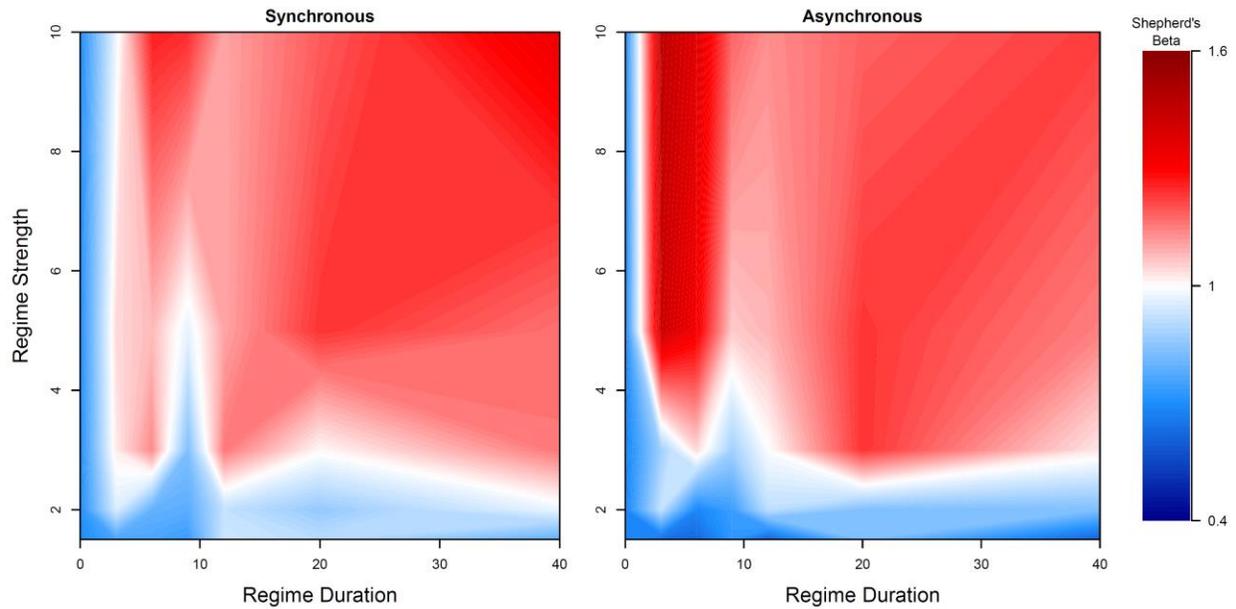


Figure 42. Median estimates of the Shepherd's β parameter, when the Shepherd spawner-recruit relationship is fit to estimates from the run-reconstruction model under different scenarios of productivity regime and sub-population synchrony. Red values indicate estimates where overcompensatory dynamics were detected, while white and blue values indicate Beverton-Holt compensation or weak compensation being detected. Values of regime strength indicate how many times greater the recruits-per-spawner at the origin is during high productivity periods is than during low productivity periods (e.g., 4 is 4-fold above the low productivity recruits-per-spawner at the origin of 3). The regime duration indicates the average duration in years of each high or low productivity regime.

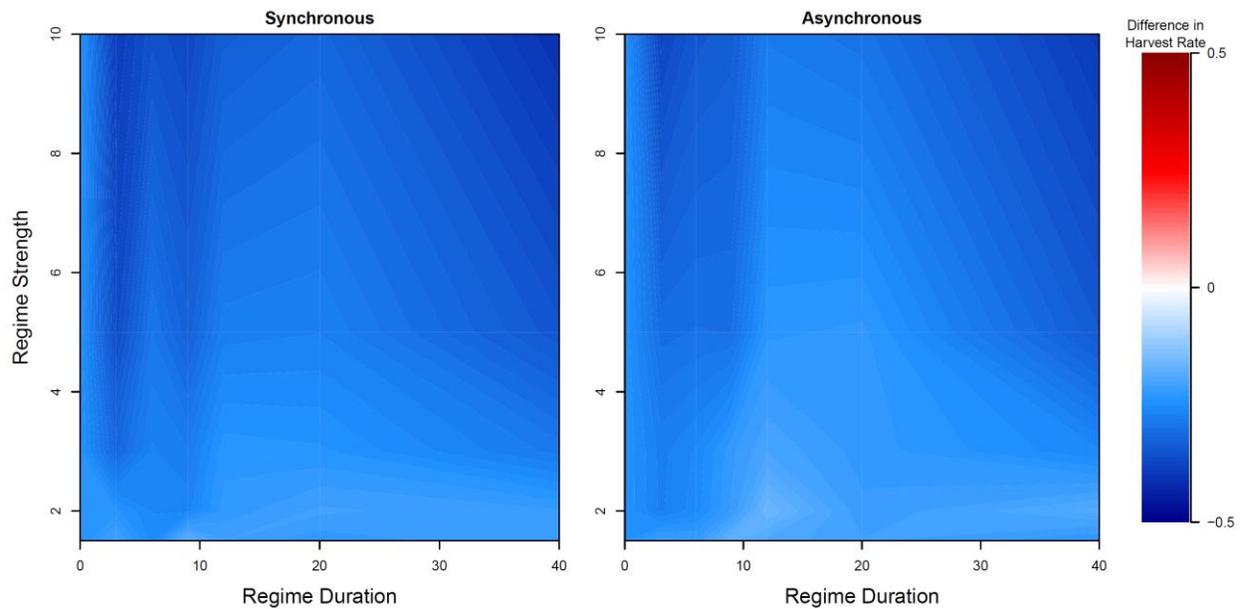


Figure 43. Median estimates of the difference in harvest rate that would produce MSY derived from a Ricker production function fit to run reconstruction estimates, compared to that which produced the maximum yield in simulated projections (i.e., estimated harvest rate at MSY minus the harvest rate that produced the maximum yield in simulated projections). Blue (red) values indicate that the estimate from the Ricker model was less (greater) than that which would produce the maximum yield. Values of regime strength indicate how many times greater the recruits-per-spawner at the origin is during high productivity periods is than during low productivity periods (e.g., 4 is 4-fold above the low productivity recruits-per-spawner at the origin of 3). The regime duration indicates the average duration in years of each high or low productivity regime.

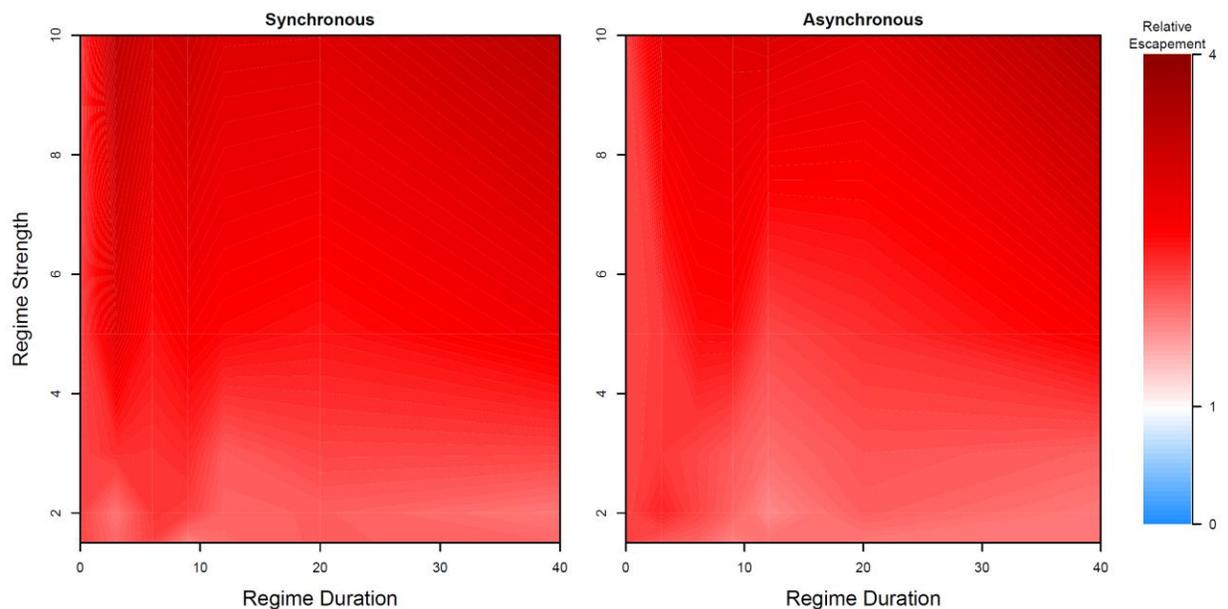


Figure 44. Median estimates of the spawning escapement that would produce MSY derived from a Ricker production function fit to run reconstruction estimates divided by the optimal escapement, i.e, the one that produced the maximum yield in simulated time-series projections. Red (blue) values indicate that the estimate from the Ricker model was greater (less) than that which would produce the maximum yield. Values of regime strength indicate how many times greater the recruits-per-spawner at the origin is during high productivity periods is than during low productivity periods (e.g., 4 is 4-fold above the low productivity recruits-per-spawner at the origin of 3). The regime duration indicates the average duration in years of each high or low productivity regime.

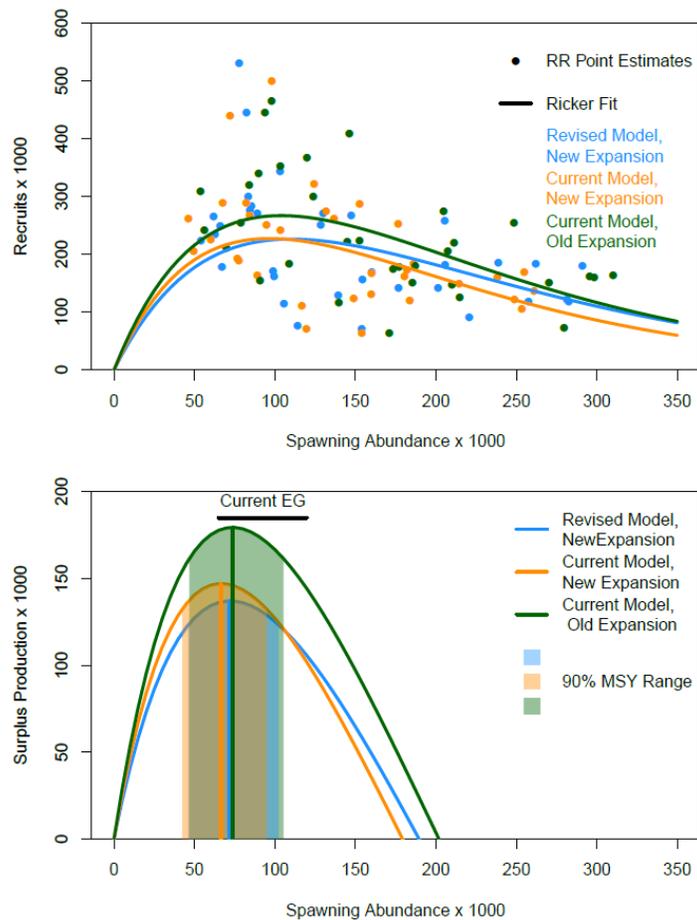


Figure 45. (Top) Comparison of relationship between spawning stock size and the subsequent recruitment for Chinook salmon in the Kuskokwim River, based on output from three different run-reconstruction (RR) models. Data points include estimates from the original model and the old lower river expansion factors from Schaberg et al. (2012) (dark green), the original model and the new lower river expansions (orange), and with a revised model using the new lower river expansions (blue). Here, the revised model includes log-normally distributed errors on weir and aerial survey indices, pooled variances for each of these survey types, and a reformulated harvest likelihood that allows model convergence. For analyses using the original model, we show output from the solution with the lowest negative log-likelihood here. Curves show best fits of the Ricker stock-recruit function fit with log-normally distributed and autocorrelated errors. (Bottom) Curves showing expected surplus production as a function of escapement for each of the Ricker functions shown in top panel. Vertical lines show S_{msy} , the escapement that would produce MSY at equilibrium. The shaded regions show the range of escapements that would produce yield that is at least 90% of expected MSY at equilibrium. The black horizontal bar denotes the current range of the ADFG escapement goal (“Current EG”).

Appendix 1. Original scope of work from the Steering Committee of the Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative (AYK SSI), commissioning this report.



To: AYK SSI Steering Committee

From: STC Sub-Committee: Milo Adkison, Daniel Schindler, Andrew Munro and John Sky Starkey (AYK SSI Steering Committee Representative)

Date: May 22, 2016

Subject: Scope of Work for STC Recommended Project: Independent Peer Review Panel to Review Run Reconstruction & Spawner-Recruit Models for Chinook salmon in the Kuskokwim River

OVERVIEW:

At their April 15, 2016 meeting, the SC took two actions with regard to the STC proposed project “Independent Peer Review Panel to Review Run Reconstruction & Spawner-Recruit Models for Chinook salmon in the Kuskokwim River.”

First, the SC asked Sky Starkey to serve as a SC rep. to the existing STC Subcommittee, which consists of Milo Adkison, Daniel Schindler and Eric Volk (now replaced by Andrew Munro from ADF&G). Second, the SC requested the Subcommittee to come back to them in 4-6 weeks with a draft project scope of work – presented below- addressing the follow two components:

1. Initial phase wherein an Independent Peer Review Panel would conduct a review and evaluation of the structure and performance of the drainage-wide run reconstruction and spawner-recruit models for generating estimates of stock status, trends, and productivity.
2. Second phase wherein the Independent Peer Review Panel would meet and collaborate with ADF&G's modeling team to discuss the findings of the panel’s review and cooperatively explore approaches to integrating panel recommendations into future modeling efforts.

INTRODUCTION:

The run reconstruction model is the core stock assessment tool to estimate annual Chinook salmon run abundance in the Kuskokwim River. The model uses observed data from weirs, aerial surveys, harvest,

and total abundance estimates based on a mark-recapture study to estimate a historical run abundance time series from 1976-2015. The model was developed by Brian Bue with input from ADF&G stock assessment scientists and was completed in 2012 (Bue et al. 2012). The original report can be accessed at: <http://www.adfg.alaska.gov/FedAidpdfs/FDS12-49>. There have been several minor changes to that model in 2014, as noted in a memorandum from Hamazaki and Liller (2015). Estimates of historic run abundance is critical component for spawner-recruit analysis and setting optimal biological escapement goals.

Effective conservation and management of salmon stocks in data-limited situations such as the AYK region require both: 1) reliable and transparent models for assessing stock status, trends, and productivity, and the associated uncertainties associated with these estimated parameters; and 2) a high degree of confidence in these models and their outputs by stakeholders and partner management agencies.

Stock assessment models informing management decisions that do not engender broad confidence inevitably generate conflict, which is costly to management agencies and stakeholders. Application of models to steeply declined stocks or stocks in rebuilding mode requires additional attention and vigilance to ensure unbiased model assessments of stock status.

Peer review is one of the foundations of science. By far the most widely accepted and robust approach to technical review and verification of stock assessment models informing management decisions – the gold standard for ensuring sound scientific advice – is use of independent expert peer review panels (NMFS 1999, NMFS 2001, NRC 1998, NRC 2002, NRC 2004). ADF&G routinely engages in other external reviews of technical analyses or stock assessments, as we are proposing here for Chinook salmon in the Kuskokwim River.

There are several reasons for proposing to conduct an independent peer review at this time:

1. Recent annual estimates of Kuskokwim River Chinook salmon run size were based on an analysis by Bue et al. (2012), who reconstructed historical sizes from several different sources of data collected from Kuskokwim River fisheries and a variety of escapement monitoring projects. This model was re-coded in 2014 to include some minor changes and was more recently re-coded in AD Model Builder with some important changes to the model structure (Martell 2016). In the process of developing this run reconstruction model – from initial development, subsequent revisions to identifying an escapement goal range (2013) for Alaska’s largest Chinook salmon subsistence fishery – it has not received a formal, independent peer review.
2. The AYK SSI Chinook Salmon Action Plan states that: “The dominant influence of density-dependent factors on recruitment dynamics in Kuskokwim River Chinook salmon is anomalous when compared with the other stocks analyzed by Catalano (2012).” The apparently strong density-dependence suggests that escapement goals should be near the lower end of historically-observed spawning abundance, which is risky unless there is high confidence in the analysis. Within a precautionary framework, the combined presence of this anomalous spawner-recruit relationship with steep declines and on-going conservation concerns warrants formal external review. The STC Subcommittee is concerned that several recent analyses using a similar run reconstruction or escapement goal methodology have suggested escapement goals near or below the lowest historical escapement.
3. These models have implications for other salmon stocks. There are plans and initial efforts to apply a related model to Chinook salmon returning to the Canadian portion of the Yukon River. We can reasonably expect stakeholders in the Yukon River watershed to expect answers to

these same questions regarding the structure and performance of the model. It would be desirable to complete a more comprehensive evaluation of the structure and performance of these run reconstructions prior to applying them to other data-limited cases in the AYK region and beyond.

4. These run reconstruction models have implications for management of other species – namely impacts on the total allowable catch of Bering Sea pollock – and are now inextricably linked to the NMFS and the NPFMC process via the “three river index” and the Proposed Rule for Chinook salmon bycatch. The NMFS has a robust external technical review for stock assessments conducted by the NPFMC BSAI and GOA Plan Teams and the Council’s Scientific and Statistical Committee (SSC). In addition to the Martell (2016) technical comments to the NPFMC on behalf of the Pollock fishing industry, we can reasonably expect other stakeholders involved in the Council process to press for application of these same technical review standards to the Kuskokwim run reconstruction models.
5. It is critical to appropriately assess the nature and cost of uncertainty and risk in data-limited stock assessments (NRC 1998). Evaluation of the application of a Bayesian approach to an integrated run reconstruction – spawner-recruit analysis has a number of benefits including “an enhanced ability to incorporate auxiliary information, convenient and rigorous consideration of measurement error and missing data, and a more complete assessment of uncertainty” (Fleischman et al. 2013: 401).
6. The STC Subcommittee is concerned that current assessments of the relationship between spawning stock abundance and production do not appropriately account for errors (precision and bias) in data collection and in model analyses of these data. The current relationship between spawning stock abundance (i.e. escapement) and production is suspiciously precise given the vast region encompassed by the Kuskokwim watershed and the ability to effectively monitor basin scale escapement and harvest.
7. Lastly, the timing for beginning the peer review process (Fall 2016) is such that data from the final year of the three year Mark/Recapture project will be available by about the time the Panel would begin its work. We are aware of no significant additional data sources available for inclusion in the model.

PROJECT OBJECTIVE:

Conduct, via an independent panel of highly qualified stock assessment experts, a comprehensive evaluation of the structure and performance of Kuskokwim River Chinook salmon drainage-wide run reconstruction and spawner-recruit models for generating reliable estimates of stock status, trends, and productivity.

METHODS:

This section provides an overview of proposed approach to achieving the project objective and is summarized in Figure 1 below.

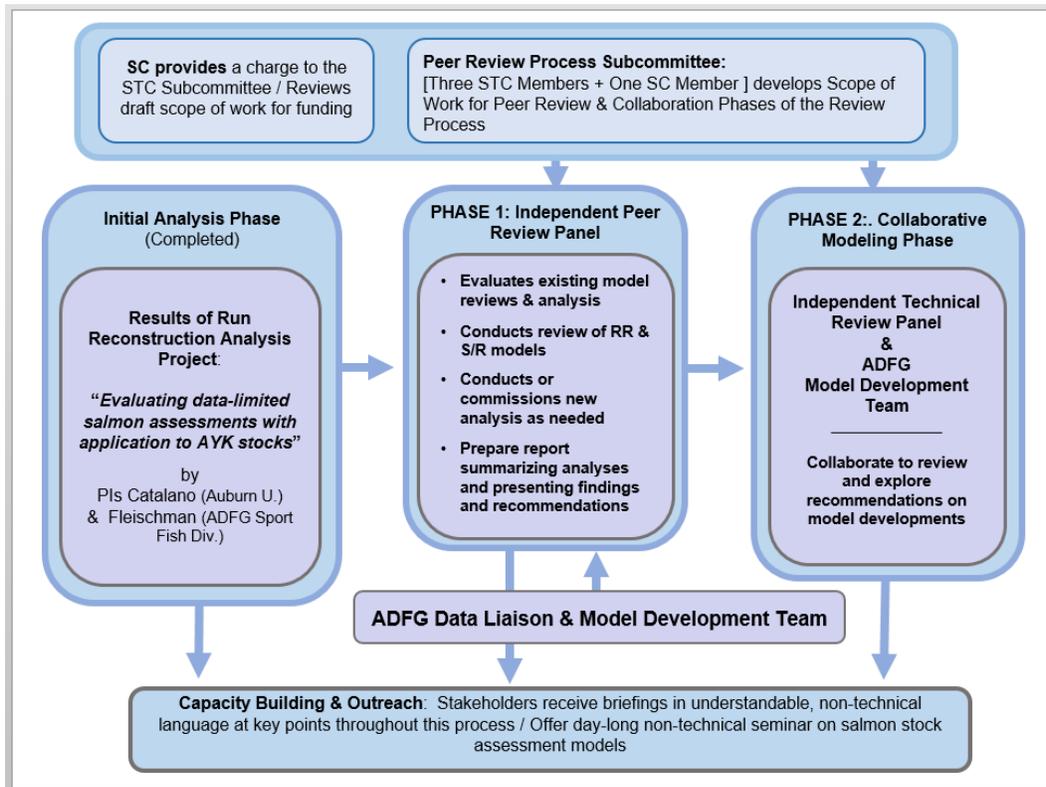


Figure 1: Overview of Independent Peer Review Process.

INITIAL ANALYSIS PHASE: (COMPLETED):

An AYK SSI funded project “Evaluating data-limited salmon assessments with application to AYK stocks” by Catalano (Auburn University) and Fleischman (ADF&G) is a first step in contributing to this broader review process by conducting three sensitivity analyses to evaluate the structure and performance of these models (See project abstract and research question in Appendix #1).

1. The first set of analyses investigated the sensitivity of the 2014 version of the Bue et al. (2012) run reconstruction (“the R model”) to five different data weighting assumptions. That report, titled “Description of the Kuskokwim River Chinook Salmon Run Reconstruction and an Investigation of Data Weighting” was prepared at the request of the KSMWG and was released summer 2015.
2. The second analysis involved challenging the Bue et al. (2012) model with simulated data as a means of exploring the sensitivity of escapement estimates to changes in simulated (i.e. specified) values of a set of stock assessment parameters. This analysis also explored a range of different data configurations that included varying the proportion of years that received a telemetric mark-recapture study, the relative proportions of aerial surveys and weirs, and the total number of escapement monitoring projects.

3. The third analysis assessed the sensitivity of the integrated run reconstruction spawner-recruit model to changes in structural assumptions of the run reconstruction model.

The latter two analyses will be available in their project final report to AYK SSI in late May, 2016

PHASE 1: INDEPENDENT PEER REVIEW PANEL PROCESS:

Appointment of Independent Peer Review Panel members: Four experts with a range of appropriate stock assessment modeling expertise will be appointed to the Independent Peer Review Panel by the non-conflicted members of the AYK STC. To maintain independence of the panel, and avoid conflicts of interest, panel members must agree to abide by the “Conflict of Interest Policy for AYK SSI Peer Reviewers” (Appendix #2). Academic and private sector members of the Panel will be compensated for their work on the panel.

The Independent Peer Review Panel will request ADF&G to identify a data and model liaison to work with the Panel (propose STC member Andrew Munro to serve in this capacity). As with other AYK SSI expert panel processes (e.g., The Escapement Goal and Chinook Action Plan Expert Panels) the AYK SSI staff Research Coordinator and Program Manager are available to assist the panel Chair and members in implementing this project.

The Panel will begin their review with thorough consultation with the ADF&G data liaison and the relevant modelers from ADF&G familiar with the data and development of the models in order to:

1. Access all relevant data, metadata, and model code.
2. Ensure a thorough understanding the data sources, their limitations, and structure of the models.
3. Ensure a thorough understanding of any anticipated revisions to the models.

Second, the Panel will review and consider the findings and sensitivity analyses presented in the Catalano & Fleischman project final report as discussed in the initial model review phase described above.

Third, building on the findings of existing model reviews, the Panel will identify and commission additional analyses, as needed, to further explore topics regarding model structure and performance as well as the limitations and uncertainties of historical stock assessment data, and new data that might become available under an expanded research program. The Panel would identify independent expert biometricians to conduct these analyses under their direction.

Potential Model Analyses / Review Topics:

The entire set of analyses and model review topics cannot be determined in advance of convening the Panel. However, we provide below a summary of some potential model analyses and review topics:

- The primary objective of the analyses will be to evaluate the reliability and sensitivity of the model outputs, given uncertainties in data and in model formulations, that most affect management. These include estimates of optimal escapement goal ranges and forecasts of next season’s returns.
- The recommended approach is an evaluation similar to that employed by Catalano, Staton, and collaborators, who simulated stock dynamics, data collection, and data analysis to compare estimation outputs to the simulated true stock properties.
- The scenarios explored will be refined in the course of further meetings and examination of preliminary results. However, the STC expects the following types of scenarios to be investigated:

1. Assessment of relationships between escapement and production at the sub-basin scale where data are the high quality (i.e. tributaries with weirs), compared to basin-wide assessments of these relationships
2. Low contrast in historical escapements; either low escapements from historically high harvest rates, or high escapements due to historically low harvest rates
3. Alternative assessment model assumptions; examining varying degrees of asynchrony in substock components, both on a year-to-year basis or in longer-term trends
4. Regime shifts resulting in persistent change in stock productivity or escapement quality both at the river scale and at the sub-basin scale
5. Loss of or degradation in the quality or quantity of data of different types, as well as improvement in data quality or the addition of new types of data (e.g., new adult abundance estimates).

It is anticipated that during this phase, the Panel will regularly consult with the ADF&G modeling team regarding planned analysis and questions about model data sources.

Following completion of their analyses and review, the Panel will prepare a technical report presenting their findings and any recommendations for improvements. This will include an executive summary appropriate to non-technical audiences.

PHASE 2: COLLABORATIVE MODELING PHASE

Independent Peer Review Panel findings are not binding on ADF&G. However, they are intended to provide expert scientific advice about strengths of existing models, potential improvements to model performance in response to particular types of data or sources of uncertainty or insights about the best approaches to incorporate new data into run reconstruction and spawner-recruit models.

This “Collaborative Modeling Phase” is designed to provide time and resources to support meetings between Panel Members and the ADF&G modeling team to discuss findings and collaboratively explore and implement recommendations on model development or data collection that may be helpful to ADF&G. This phase may include a mix of more formal collaborative discussions and informal technical sessions where participants are working directly with data sets and model code to explore alternate approaches.

DELIVERABLE PRODUCTS:

1. Report and Recommendations from the Independent Peer Review Panel (Dec. 2017)
2. Capacity Building / Outreach Seminar (Fall 2017)
3. Summary of post-review collaborative model session with peer review panel members and ADF&G Modeling Team.

PEER REVIEW PANEL PROCESS TIMELINE:

| TASKS | 2016 Jun-Dec | 2017 Jan-June | 2017 July-Dec | 2018 Jan- Jun |
|--|---|---------------|---|---|
| Appoint Independent Peer Review Panel | Peer Review Panel Appointed by Sept.1 | | | |
| Access Code & Data | Work with ADF&G Data Liaison; initial consultation session with ADF&G modeling team. | | | |
| Conduct Analysis and Peer Review | Independent Peer Review Phase: Sept. 2016-Dec. 2017. Includes three in-person meetings during the course of this phase. | | | |
| Capacity Building Seminar | Periodic briefings in non-technical language at key points during this process. | | | Outreach / Capacity Building seminar: Jan. or Feb.2018 |
| Final report | | | Draft Final report produced by Nov. 15 th ; finalized by Dec. 15 th | |
| Collaborative Modeling Phase | | | | Two in-person meetings / collaborative modeling sessions with Panel & ADFG modeling team (Jan & March 2018) |

CAPACITY BUILDING:

- Stakeholders will receive periodic briefings in understandable, non-technical language at key points throughout this process.
- As the Peer Review Panel completes their work on Phase 1, we propose to present day-long non-technical seminar on salmon stock assessment models for interested stakeholders.

REFERENCES:

Bue, B. G., K. L. Schaberg, Z. W. Liller, and D. B. Molyneaux. 2012. Estimates of the historic run and escapement for the Chinook salmon stock returning to the Kuskokwim River, 1976–2011. Alaska Department of Fish and Game, Fishery Data Series No. 12-49, Anchorage.

Catalano, M.J. 2012. A Comparative Analysis of Patterns in Productivity for AYK and Selected Non-AYK Chinook Salmon Stocks. Workshop Background Paper (unpublished working paper). *In* AYK SSI Chinook Salmon Synthesis Workshop, May 2-3, 2012. Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative, Anchorage, AK.

Fleischman, S.J., M.J. Catalano, R.A. Clark, and D.R. Bernard. 2013. An age-structured state-space stock-recruit model for Pacific salmon (*Oncorhynchus* spp.). *Canadian Journal of Fisheries and Aquatic Sciences* 70:401-414.

Martell, S. 2012. Technical Comments on submitted March 8, 2016 to the NPFMC during the public comment period for the implementation of Bering Sea Chinook salmon bycatch measures for the Bering Sea pollock fishery per NPFMC amendment 110.

<https://www.regulations.gov/#!documentDetail;D=NOAA-NMFS-2015-0081-0019>)

National Marine Fisheries Service. 1999. *National Marine Fisheries Service: Ecosystem-Based Fishery Management*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration. NOAA Technical Memorandum NMFS-F/SPO-33.

National Marine Fisheries Service. 2001. *Marine Fisheries Stock Assessment Improvement Plan*. Report of the National Marine Fisheries Service National Task Force for Improving Fish Stock Assessments. U.S. Department of Commerce, National Oceanic and Atmospheric Administration. NOAA Technical Memorandum NMFS-F/SPO-56.

National Research Council. 1998. *Improving Fish Stock Assessments*. National Academy Press, Washington, D.C.

National Research Council. 2002. *Science and Its Role in the National Marine Fisheries Service*. National Academies Press, Washington, D.C.

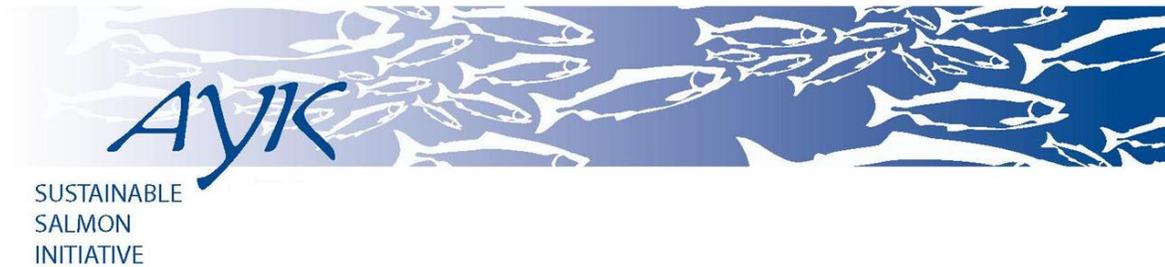
National Research Council. 2004. *Improving the Use of the "Best Scientific Information Available" Standard in Fisheries Management*. National Academy Press, Washington, D.C.

Appendix 1a: Catalano & Fleischman Funded AYK SSI Project Abstract & Research questions:

| | |
|-------------------------|--|
| Project Title: | Evaluating data-limited salmon assessments with application to AYK stocks |
| Investigator(s): | Dr. Matthew J. Catalano, Principal Investigator, Assistant Professor, Department of Fisheries and Allied Aquacultures, Auburn University Steve Fleischman, Fisheries Scientist, Division of Sport Fish, Alaska Department of Fish and Game, |
| Project Period: | October 1, 2013 – February 29, 2016 |
| Study Location: | Kuskokwim River and other large western Alaska drainages |
| Abstract: | <p>Evaluating hypotheses for AYK Chinook salmon (<i>Oncorhynchus tshawytscha</i>) declines and managing these stocks requires that we obtain reliable estimates of stock status, trends, and productivity. These estimates have been obtained from novel drainage-wide run reconstruction models developed specifically for data-limited cases. However, these models have potential pitfalls that warrant further investigation. In particular, the models may be susceptible to bias stemming from incomplete temporal and spatial coverage of run enumeration projects and natural between-stock variation in population dynamics within a drainage. We will investigate the circumstances under which these drainage-wide reconstructions provide accurate and precise estimates of Chinook salmon abundance and productivity. We undertake this work from a Bayesian perspective, which means that we will incorporate into the models different types of uncertainty and will allow the inclusion of prior information on stock abundance and productivity. First we will use a series of computer simulation-estimation analyses to test the influence of (1) among-stock within-drainage temporal variation in population dynamics, (2) data quality, (3) data quantity, and (4) data type on the accuracy and precision of these models. We will also consider the costs of data collection scenarios in light of model performance. This analysis will provide guidance on the reliability of these models under different data collection scenarios and their associated costs, which will aid prioritization of field sampling programs. Second, we will apply the Bayesian state-space formulation of the model to the Kuskokwim River Chinook salmon stock and conduct a thorough sensitivity analysis of abundance, productivity, and hypothetical escapement goals. The application of the Bayesian approach to this stock will more thoroughly deal with the many uncertainties in this system. The Kuskokwim run reconstruction relies on drainage-wide estimates of total abundance from telemetric mark-recapture and empirical abundance-drainage area models. We will explore new methods for salmon telemetric mark-recapture analyses and will conduct an analysis of associations between Alaskan Chinook salmon abundance accessible drainage area. This project will provide a starting point for any future efforts to incorporate risk and uncertainty into harvest policy analyses for the Kuskokwim Chinook salmon stock.</p> |

Objectives and Project Design

- 1. Rationale & Project Research Question:** We will investigate whether, and under what circumstances, drainage-wide Bayesian state-space run reconstruction models provide accurate and precise estimates of Chinook salmon abundance, productivity, and management quantities. These models are increasingly being used to assess AYK Chinook salmon stocks, but the models have not been fully tested for performance under realistic ranges of population dynamics parameters, data quality, quantity, and type (e.g., weir, air survey, mark-recapture). We will use a simulated hypothetical drainage that emulates the Kuskokwim River watershed to evaluate how variation in population dynamics among stocks within the drainage, as well as data quantity, quality, and type affect the accuracy, precision, and cost of estimates of abundance and productivity from these models. We hypothesize that performance of these models will be reduced when temporal variation in productivity is weakly correlated among stocks within the drainage, and under low data quality, and quantity. We will then apply the model to historical data from the Kuskokwim River stock and compare the results with the Bue et al. (2012) penalized maximum likelihood approach. We hypothesize that the application of the Bayesian approach to the Kuskokwim Chinook case will result in greater uncertainty in model estimates, which will carry through to estimates of management quantities such as escapement that produces maximum sustained yield (Smsy). This more thorough treatment of uncertainty could benefit future policy setting efforts by facilitating structured thinking regarding the risks associated with various harvest policies. Development of an alternative analysis for total abundance estimates from mark-recapture and abundance-area relationships will improve management by potentially reducing bias and uncertainty in these estimates.



CONFLICT OF INTEREST POLICY FOR AYK SSI PEER REVIEWERS

As a peer reviewer you have been asked to review a research proposal for funding by the AYK Sustainable Salmon Initiative. The performance of your review requires that you be aware of potential conflicts of interest. Prior to reviewing this proposal we ask that you read the “Conflict of Interest Statement for Peer Reviewers” below and review the examples of potentially biasing affiliations or relationships listed below. Please complete the certification statement below if conflicts of interest do not exist and return via e-mail with your completed review or fax it to the AYK SSI Office at 907-258-6688. If you cannot sign this form due to a conflict of interest, please contact us immediately and return all materials to the AYK Sustainable Salmon Initiative. Conflicts of interest are not accusations and do not imply that a reviewer’s judgment is compromised.

Conflict of Interest: A conflict of interest exists when financial interests, or other opportunities for tangible personal or professional benefit, could be construed to exert an improper influence on a peer reviewer’s professional judgment. Improper influence could be used not only to advance one’s own research program but also to promote unfairly a colleague’s or former graduate student’s program. Thus, care must be exercised to avoid a conflict of interest during the peer-review process and the discussion of past, current, or proposed research. A conflict of interest also exists when circumstances simply have the appearance of compromising the professional judgment of a member. Conflicts of interest are not accusations and do not imply that a member’s judgment was compromised. It is the policy of the AYK Sustainable Salmon Initiative’s Scientific and Technical Committee that conflicts of interest should be avoided wherever possible and disclosed and minimized in situations where interests cannot be reasonably separated.

Confidentiality: The AYK Sustainable Salmon Initiative receives proposals and manuscripts in confidence and protects the confidentiality of their contents. Everything about the proposal must be kept in strict confidence. If, as a peer reviewer, you gain access to information not generally available to the public, you must not use that information for your benefit or make it available for the benefit of any other individual or organization without the permission of the authors. You must not quote or otherwise disclose or use material from any proposal that you review. You are not to discuss proposals or manuscripts with its authors or other colleagues. Questions about the proposal are to be discussed with the person coordinating the review of this proposal but with no one else.

Potential Conflicts of Interest for AYK Sustainable Salmon Initiative Peer Reviewers

Relationship with the principle investigator(s).

1. Known family or marriage relationship.
2. Business or professional partnership.
3. Past or present relationship as a thesis advisor or thesis student.
4. Other relationships, such as close personal friendship, that might tend to affect your judgment or may be seen as doing so by a reasonable person familiar with the relationship.

Relationship to the principle investigator(s) institution:

1. Current employment at the institution associated with the proposal; or employment with the institution via consulting or an advisory arrangement; or employment within the last 12 months, or you are currently being considered for employment with the institution.
2. Holder of any office, governing board membership, or relevant committee chair in the institution.
3. Current enrollment as a student in the department or school of the institution that originates the proposal.

Confidentiality of Peer Reviews and Reviewer Identities

AYK Sustainable Salmon Initiative policy is that reviews and review identities will not be disclosed except that verbatim copies of reviews (without name and affiliation of the reviewer) will be sent by staff to the principal investigator.

Appendix 2. Review of ADFG methodology used in mark-recapture studies to estimate abundance of Chinook salmon in the Kuskokwim River; by Dr. John Skalski, University of Washington.

COLUMBIA BASIN RESEARCH

Review of the Mark–Recapture Studies to Estimate Adult Chinook Salmon Abundance in the Kuskokwim River, 2003–2007, 2014–2016

TO: ARCTIC–YUKON–KUSKOKWIM REGION - SUSTAINABLE SALMON INITIATIVE

431 West 7th Avenue, Suite 204, Anchorage, Alaska 99501

FROM: JOHN R. SKALSKI

Columbia Basin Research, School of Aquatic and Fishery Sciences, University of Washington
1325 Fourth Avenue, Suite 1515, Seattle, Washington 98101-2540



Table of Contents

| | |
|--|-----|
| Overview | 101 |
| Response to Review Questions..... | 101 |
| Review of Statistical Methods to Estimate Kuskokwim Chinook Salmon Abundance | 104 |
| 2014–2016 | 104 |
| 2002–2007 | 107 |
| Literature Cited | 108 |
| Appendix A: Var(\hat{N}) in Stages | 109 |
| Appendix B: Variance of m, the Estimated Number of Mark Recaptures | 112 |

Overview

This review of the mark–recapture estimates of Chinook salmon abundance in the Kuskokwim River is based on the annual reports for 2014–2016 by Head et al. (2017) and Smith and Liller (2017a, b) and the compilation for 2002–2007 by Schaberg et al. (2012). The review examined the assumptions, estimation, and variance calculations for the single mark–recapture studies performed in those years. No attempt was made to independently check the numerical calculations for accuracy.

The mark–recapture studies are based on the joint use of dual radio/spaghetti-tagged and spaghetti-tagged adult Chinook salmon. Considerable effort and diligence went into the design and conduct of the studies that involved tracking the fate of the radio-tagged fish over vast areas and the collection of recovery data from multiple weirs. Attention to model assumptions and their implications to mark–release and recovery data collection are evident. The goal of estimating annual abundance within a $\pm 25\%$ (i.e., CV = 12.5%) were commonly met, 2014–2016.

This review is organized starting with a summarial response to technical questions that were proposed for the review. The next section is a technical overview of the mark–recapture reports, followed by specific comments, and concluding with analytical appendices. The review was so structured that interested parties can read the review at different technical levels of specificity as the report progresses.

Response to Review Questions

Question 1: Are the current mark-recapture methods used to estimate abundance of Kuskokwim River Chinook salmon likely to produce unbiased estimates?

The mark–recapture studies have been refined over time as the understanding of how to sample the system improved and as concerns over model assumptions have been elevated. The 2014–2016 studies are well crafted and the model assumptions are taken seriously. The reanalysis of the 2002–2007 data is in light of this improved knowledge.

The single mark–recapture model requires at least one of the two samples be representative of the population. In a multi-run, multiple-recapture-location scenario, as is the case with the Kuskokwim River, this assumption can be easily violated if sufficient care is not taken. The study has taken great lengths, both logistically and analytically, in recent years to help assure this assumption is met. The 2002–2007 retrospective reanalysis evaluated this assumption in 2007 and applied the results to prior years. The study also correctly inferred from the Lincoln/Petersen index when mortality only is operating in the population, the abundance is being estimated at the time (or place) of the initial sampling. Tag loss based on dual-tagging of fish is also being properly estimated and taken into account.

Other model violations such as post-release handling effects and high grading of the mark–released fish could be sources of bias of unknown size. The unexpected number of radio-tagged fish that residualized in the mainstem without entry into the natal tributaries could be indicative of delayed handling effects that should be further investigated. Irrespective of this comment, I believe no major sources of bias exist in these studies.

Question 2: What are the key assumptions of the current mark-recapture experiments and are they likely to be met given the current experimental design and implementation?

The key assumptions of the model are the random and independent sampling of the population in the two periods of the Lincoln/Petersen index. This assumption can be relaxed to at least one of the two periods representatively sampling the population. The investigators correctly identified this concern when estimating the abundance in the complex system of runs and natal tributaries within the Kuskokwim River. These assumptions have been met with the implemented logistics of the study and thoroughly tested with the data collected.

Other key assumptions include the number of fish marked (n_1) and examined for marks (n_2) are known without error and marked–recaptured fish (m) are properly identified. The assumption of n_1 , n_2 , and m known without error can be relaxed to allow the quantities to be estimated unbiasedly, as long as the variance calculations correctly take into account these extra sources of sampling error. The total weir passage (n_2) is not enumerated but must be estimated (\hat{n}_2) for outages, etc. The reports indicated as much as 99% of the weir passage is enumerated and only small adjustments are used. The exact method of adjustment and the size of adjustments are not specified, and that source of uncertainty is currently not incorporated into the error variance for \hat{N} . The number of radio-tagged, mark-recaptured fish (m) is adjusted for tag loss by an appropriate approach, assuming the loss rate for spaghetti tags is zero.

These key assumptions appear to be met or properly accounted for by analytical adjustments to the abundance estimator.

Question 3: Are the levels of precision currently reported from mark-recapture experiments likely to appropriately reflect the uncertainties in estimates of abundance, given the possibility of violating key assumptions described above?

The investigators have chosen to use Monte Carlo bootstrap techniques to estimate the variances and construct 95% confidence intervals (CIs) for the annual abundance estimates (\hat{N}). These techniques are appropriate if the sampling processes are properly simulated and all sources of variance incorporated. The sampling scheme used in the Kuskokwim River is rather simple, and the variance of \hat{N} can be derived analytically to produce a closed-form expression for the variance of \hat{N} that explicitly incorporates all sources of variance (Appendices A and B). This variance formula should be used to determine where the Monte Carlo variance estimates comport. The advantage of the Monte Carlo bootstrap estimates if done correctly is in CI construction. At small sample sizes, \hat{N} may not be normally distributed and the Monte Carlo CI estimates will capture the right-hand skewed distribution. In variance calculations, I recommend 10,000, not 1,000, simulations to accurately estimate the variance. The advantage of the analytical variance is the ability to assess the relative contributions of the error sources to overall precision.

The description of the bootstrap procedure is incomplete and vague, making it difficult to assess whether the sampling processes are being modeled correctly. Generation of m 's is a compound process that is not being explicitly described. The variance in the estimate of the weir passage counts (\hat{n}_2) is not being incorporated in the variance calculations. On the other hand, the number of marked releases (n_1) is a known constant for radio-tagged releases but is being treated as a binomial random variable unnecessarily.

The largest contributor to the variance of \hat{N} is the mark–recapture process. Contributions associated with the estimation of \hat{m} , \hat{n}_2 , and possibly \hat{n}_1 , are relatively smaller. A quick check suggests the report coefficients of variation (CVs) are of the proper orders of magnitude but attention to detail will improve the confidence in the reported results.

Despite a statement made in the 2015 (p. 12) annual report, the Monte Carlo methods estimate the variance of \hat{N} only under the assumption of the nominal mark–recapture model. The simulations do not capture the effect of model violations on the estimates of N_1 nor are they robust to model violations. Separate simulations under alternative model scenarios would be necessary to perform a robustness evaluation.

Note the variance of \hat{N} is $\propto \hat{N}^2$ (Equation A9). This implies $\widehat{\text{Var}}(\hat{N})$ and \hat{N} are not independent. For all else being equal, the variance estimate will increase as the abundance estimate increases. Consequently, low abundance years will look relatively more precise than larger abundance years. In performing weighted (w) analyses using the abundance estimates, it is therefore not appropriate to weight inversely proportional to the variance, i.e.,

$$w_i \propto \frac{1}{\widehat{\text{Var}}(\hat{N}_i)}$$

but instead to weight

$$w_i \propto \frac{1}{\left(\frac{\widehat{\text{Var}}(\hat{N})}{\hat{N}^2}\right)} = \frac{1}{\text{CV}(\hat{N})^2}.$$

In this case, $\text{CV}(\hat{N})$ is roughly independent of \hat{N} , and the weights are proportional to the relative error.

Question 4: Given that some details of the mark-recapture experiments have changed among years (e.g., location of release site), what are the appropriate ways to use data from mark-recapture experiments, in years where no such experiments were run?

The Schaberg et al. (2012) report went a long way in bringing the 2003–2007 mark-recapture surveys to the same level of rigor observed in the 2014–2016. The 2014 survey estimates abundance to rkm 294, while 2015 and 2016 estimate abundance to rkm 67.

It is not clear for the purposes of run reconstruction whether abundance estimates for the years 2008–2013 are needed. Nevertheless, if abundance values are desired for the intermediate years, perhaps a regression estimator could be used regressing the estimates of \hat{N}_1 in available years against escapement indices (I_i) based on weir counts, i.e.,

$$\hat{N}_u = \alpha + \Sigma B I_i.$$

Then using comparable weir count indices in out-years, project the abundance that might have occurred in those missing years. The projection will likely be noisy.

Question 5: How should abundances of fish returning to tributaries downstream of the release site be estimated and incorporated into estimates of total in-river abundance?

The 2015–2016 mark–recapture studies were adjusted to estimate abundance above rkm 67 instead of rkm 294 by adjusting the release point for n_1 . This seems a good strategy to reduce the amount of lower river abundance uncovered by the mark–recapture study. In the area below rkm 67, a ratio-type estimator (Schaberg et al. 2012:11, Equation 10) or regression estimator seems most appropriate. While Equation 10 of Schaberg et al. (2012) is suited for a single year, the model could be extended to a regression approach across multiple years. The multi-year regression model would more realistically estimate uncertainty because it would include lack-of-fit, which variance Equation 14 of Schaberg et al. (2012) does not. In estimating the variance for the predictor, it is also important to use the proper variance for predicting a “new y ” (Kutner et al. 2004:55–60).

Review of Statistical Methods to Estimate Kuskokwim Chinook Salmon Abundance

2014–2016

A combination of dual radio-tagged/spaghetti-tagged fish and spaghetti-tagged fish were used to estimate Chinook salmon abundance in the Kuskokwim River, Alaska, 2014–2016. While superficially similar, all three studies using a variation on the Lincoln/Petersen index two-sample, mark–recapture model, the specifics of the estimation process varied between years (Table 1). The abundance estimator is based on three presumably known quantities:

n_1 (i.e., M') = number of tag released fish available for recapture,

n_2 (i.e., C') = number of fish examined in period 2 for tags,

m (i.e., R') = number of mark–recaptured fish in period 2.

The three years of study differ with regard to which of these quantities were estimated and how they were estimated. The overall abundance estimator does not differ but the variance of the abundance estimator does. The annual reports used Monte Carlo simulation techniques to estimate the variance and 95% CIs. Analytical techniques to estimate the variance would improve the transparency of the process and would provide indications of the sizes of the different sources of sampling variance. The most important reason for estimating the CI by Monte Carlo simulation is concern for nonnormality and asymmetry of the interval estimates. The small sample size in 2014 justified the use of the Monte Carlo approach in 2014 but was unnecessary in 2015–2016 because of larger sample sizes and near symmetric interval estimate results.

Table 1. Summary of the nature of the mark–recapture quantities used in estimating Chinook salmon abundance in the Kuskokwim River, 2014–2016, along with tagging method and study areas.

| | 2014 | 2015 | 2016 |
|-------|-----------------------------|-----------------------------|-----------------------------|
| n_1 | Known | Estimated | Known |
| n_2 | Assumed known without error | Assumed known without error | Assumed known without error |

| | | | |
|--|-----------|----------------|-----------|
| m | Estimated | Estimated | Estimated |
| Primary tag used in mark–recapture ratio | Radio tag | Spaghetti tag* | Radio tag |
| Abundance estimated upstream from rkm | Rkm 294 | Rkm 67 | Rkm 67 |

* \hat{N} could have also been estimated using radio-tag recoveries

Taking the variance of the Lincoln/Petersen index in stages, an overall variance of \hat{N} , taking into account uncertainty in all 3 of the mark–recapture counts (i.e., \hat{n}_1 , \hat{n}_2 , and \hat{m}), can be written as follows:

$$\begin{aligned} \text{Var}(\hat{N}) \approx N^2 \left(\left(\frac{N}{mn_2} \right) + \left(\frac{2N}{n_1 n_2} \right)^2 + 6 \left(\frac{N}{n_1 n_2} \right)^3 \right) + N^2 \cdot \text{CV}(\hat{n}_1)^2 + N^2 \cdot \text{CV}(\hat{n}_2)^2 \\ + \left(N^2 + \text{Var}(\hat{N}) \right) \cdot \text{CV}(\hat{m})^2. \end{aligned} \quad (1)$$

Special annual cases can be readily obtained by setting the coefficient of variation (i.e., $\text{CV}(\hat{\theta}) = \text{SE}(\hat{\theta})/\theta$) to zero for quantities known without error. All of the 2014–2015 analyses assumed the weir passage counts were measured without error (i.e., $\text{CV}(\hat{n}_2) = 0$). In other years, m and sometimes m and n_1 were estimated with error (Table 1). There are indications the Monte Carlo simulations did not always identify the proper error sources to incorporate in the analyses. Their results should be checked against Equation (1) for comport.

If the mark–recapture estimates are going to be used to calibrate the reconstruction model, the annual estimates should be comparable. In 2014, the baseline of the study area was rkm 294; in 2015–2016, it was rkm 67. Using the 2015 radio-tag data with a probability of migrating from the tag site to rkm 294 of 505/623, the abundance estimate is $\hat{N} = 107,645.7$, down from the reported value of $\hat{N} = 115.541$ at rkm 67. In 2016, the abundance at rkm 294 would be estimated at $\hat{N} = 107,310.9$, compared to the reported value of 120,000 at rkm 67.

Extraordinary efforts were used to determine the fate of each radio-tagged fish and assess model assumptions. In a mixed run, multiple recovery site study, it is easy to violate the assumptions of equal capture probabilities and independent sampling. In a Lincoln/Petersen index study, at least one of the two sampling periods needs to be a representative sample of the population. The statistical methods outlined in Appendix A.1 are an appropriate approach to testing these assumptions. The attention to field sampling and analytical results suggest these model assumptions are being fulfilled. However, sample sizes of marked fish are small, and marked fractions in the individual tributaries are tiny (i.e., <0.01), so the power of these tests to detect model violations should nevertheless be evaluated.

The Lincoln/Petersen index is a closed population estimator but remains valid if the population is half open (i.e., mortality only or recruitment only). When mortality only is operating during the course of the study, the model estimates abundance at the time (or place) of the initial sampling (i.e., rkm 294 or 67). The Chapman (1951) bias correction to the Lincoln/Petersen index is based on the assumption that

sampling can be modeled by a hypergeometric model where sample counts (i.e., n_1 , n_2 , and m) are integer values and n_1 and n_2 are fixed in advance. It is arguable whether that bias adjustment is applicable at the Kuskokwim River where m and perhaps n_1 are estimated with error. In which case, the traditional estimator of Petersen (1896), $\hat{N} = n_1 n_2 / m$, may be more appropriate.

In general, the calculation of the point estimates of abundance appear to be correct. However, the practice of rounding the estimates of m to a whole number is inappropriate, given \hat{m} is really an estimate, not an observed count. For instance, in 2015, rounding m from 54.6061 to 55 changes the abundance estimate from 61,689.2 to 61,255.3—not a large discrepancy, but nevertheless unnecessary and unjustified.

While the study took extraordinary steps to assure model assumptions, systematic errors could nevertheless creep into the study and should be further considered. Among the things to consider:

1. At the recovery sites (i.e., weir), were radio towers operating at times in 2014 when the weirs were not operating? If so, the mark–recapture ratio m/n_2 would be inflated and abundance underestimated.
2. Post-release handling effects or mortality of the radio-tagged fish could bias study results. Do the higher-than-expected number of radio-tagged adults remaining in the mainstem indicate a problem?
3. In any tagging study, there is the inevitable and unavoidably possibility that the fish selected for tagging are high graded and not representative of the actual population. No one wants to deliberately tag a sick or dying fish. This type of error would underestimate abundance and, if inconsistent between years, distort annual trends.
4. The studies essentially assumed the spaghetti tags had 100% retention and only radio tags failed or were lost. Were any radio-tagged fish found without their corresponding spaghetti tag to suggest the need for a dual-tag loss correction? The 2015 report (p. 10) states the weir crews “observed” the radio and spaghetti tags. Were the radio tags checked to be operational? The mark–recapture model requires tags to be both present and operationally active.
5. The 2015 collected radio-tag data should have been comparable to the 2014/2016 studies. An alternative abundance estimate could have obtained and compared to the reported spaghetti mark–recapture result. This comparison would have provided a means to assess the robustness for the mark–recapture results.
6. Contrary to that statement in the 2015 report (pp. 11–12), the Monte Carlo results to estimate CI width do not incorporate the effect of model violations. The described Monte Carlo process is only looking at variability under the existing model assumptions. Separate robustness runs would be necessary to this end, where model assumptions are specifically violated and modeled in one or more ways.

My impression is that the point estimates of abundance are likely valid; any small bias that might exist would not distort the pattern of the annual trend. On the other hand, the methods used for variance and CI calculations are less clear and possible errors in calculation could bias the weighting in subsequent run reconstructions.

The written description of the Monte Carlo simulations to estimate CIs raises some concerns on how they are being performed. With the correct assumptions and proper sampling structures, Monte Carlo

methods can be used to provide reliable variance and CI estimates. If improperly constructed, the results can be erroneous. Consider the following:

1. The key element of the Monte Carlo simulations should be the modeling of \hat{m} as a compound process, where the true m is a hypergeometric (or \approx binomial) random variable, a function of n_1 and n_2 and, in turn, the observed \hat{m} is a function of m and the tag retention rate per 2014 or observation rate per 2015–2016. As described in the methods, it is not clear this compound process is being used. Poorly estimated CIs would not impact the point estimates but could affect weights going into the run reconstruction.
2. 2014 report (p. 10), the tag release of $n_1 = 329$ is a known quantity without error. However, in the Monte Carlo simulations, it was modeled as a binomial random variable. This is an appropriate approach *a priori* when investigating the anticipated precision of the 2014 and 2016 studies, but not *a posteriori* after a particular n_1 has been realized. However, in 2015, simulating the uncertainty in \hat{n}_1 when applied to the spaghetti tags is a necessary element of the Monte Carlo process. The difference in approach occurs because in 2014/2016, the individual radio-tagged fish constituting n_1 were known; this was not the case in 2015 with the spaghetti tags.
3. The multinomial sampling of the tags to the various tributaries (2015, p. 6) is unnecessary, given marked rates were constant across the tributaries, and the data were nevertheless pooled in the analysis. This element of the 2015–2016 Monte Carlo simulations is superfluous window dressing. What is important is the hypergeometric sampling of m , given n_1 and n_2 or approximated by binomial sampling (Seber 1982: Equation 3.3) since N is large and n_2/N is small.
4. Apparently, the total estimated passage at the weirs (n_2) in the 2014–2016 reports are treated as a known constant without error, since there is no mention of it in the variance simulations. Ignoring this error source will underestimate the variance of the mark–recapture study.
5. It is often recommended 1,000 simulations to examine for bias in point estimation, but a simulation sample size of $\geq 10,000$ for variance estimation.
6. Monte Carlo results should be compared to variance estimates based on Equation (1) to assure that the sampling is properly modeled in the simulations. In general, the first term in Equation (1) will dominate the variance calculations. The second–fourth terms of Equation (1) will contribute less but should nevertheless be included.

2002–2007

The Schaberg et al. (2012) report recalculated mark–recapture estimates of Chinook salmon abundance above rkm 297 for the years 2003–2007. The recalculation was justified based on finding the various runs within the Kuskokwim River were proportionately marked in 2006 and 2007 and then applying that assumption to the years 2003–2005. They elected not to estimate abundance for 2002 because they felt the fish were not tagged in proportion to abundance of the various runs that year.

The 2003–2007 estimates of abundance were again based on the Chapman (1951) correction for the Lincoln/Petersen index. Similar to 2014–2016, the weir counts were treated as known constants despite adjustment for periods when the weirs were inoperable. This does not impact the point estimates but will underestimate the variance of \hat{N} . In these years, 2003–2007, no attempt was made to correct for

radio-tag loss/failure. Not correcting for tag loss will result in positively biased estimates of abundance. In 2014 and 2015, the radio-tag retention rates were estimated to be 0.9706 and 0.9726, respectively.

Bootstrap estimates of precision were again calculated for the 2003–2007 mark–recapture estimates. The 95% confidence intervals (CIs) were based on estimates of the 2.5% and 97.5% order statistics (Schaberg et al. 2012, Table 6), as in 2014–2016.

As in subsequent years, n_1 was unnecessarily and incorrectly treated as a binomial random variable. In the Lincoln/Petersen index based on the hypergeometric model, n_1 is assumed to be a fixed constant. In the radiotelemetry studies, the individuals constituting the n_1 release are known individually and exactly. This unnecessary additional would inflate the variance estimate of \hat{N} and brings into question whether the actual sampling process is properly understood. See other comments related to variance estimation for the years 2014–2016.

Literature Cited

- Chapman, D. G. 1951. Some properties of the hypergeometric distribution with applications to zoological censuses. University of California Publications in Statistics 1:131-180.
- Kutner, M. H., C. J. Nachtsheim and J. Neter. 2004. Applied linear regression models, Fourth edition. McGraw-Hill/Irwin, New York, NY, USA.
- Petersen, C. G. J. 1896. The yearly immigration of young plaice into the Limfjord from the German Sea. Report of the Danish Biological Station 6:1-77.
- Schaberg, K. L., Z. W. Liller, D. B. Molyneaux, B. G. Bue, and L. Stuby. 2012. Estimates of total annual return of Chinook salmon to the Kuskokwim River, 2002–2007. Alaska Department of Fish and Game, Fishery Data Series, No. 12-35, Anchorage.
- Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. Macmillan, New York.

Appendix A: $\text{Var}(\hat{N})$ in Stages

Considering the most general case for the Lincoln/Petersen index, where

$$\hat{N} = \frac{\hat{n}_1 \hat{n}_2}{\hat{m}} \quad (\text{A1})$$

and where

\hat{n}_1 = estimated number of marked animals released (n_1),

\hat{n}_2 = estimated number of animals examined for marks in second period (n_2),

\hat{m} = estimated number of mark recaptures (m).

The $\text{Var}(\hat{N})$ in this case can be derived in stages, where

$$\begin{aligned} \text{Var}\left(\frac{\hat{n}_1 \hat{n}_2}{\hat{m}}\right) &= \text{Var}_{S_4} \left[E_{S_3} \left[E_{S_2} \left[E_{S_1} (\hat{N} | S_2, S_3, S_4) \right] \right] \right] \\ &\quad + E_{S_4} \left[\text{Var}_{S_3} \left[E_{S_2} \left[E_{S_1} (\hat{N} | S_2, S_3, S_4) \right] \right] \right] \\ &\quad + E_{S_4} \left[E_{S_3} \left[\text{Var}_{S_2} \left[E_{S_1} (\hat{N} | S_2, S_3, S_4) \right] \right] \right] \\ &\quad + E_{S_4} \left[E_{S_3} \left[E_{S_2} \left[\text{Var}_{S_1} (\hat{N} | S_2, S_3, S_4) \right] \right] \right] \end{aligned} \quad (\text{A2})$$

and where

S_1 = denotes estimation of \hat{n}_1 ,

S_2 = denotes estimation of \hat{n}_2 ,

S_3 = denotes estimation of \hat{m} ,

S_4 = denotes the mark–recapture process.

The first term of Equation (A2) reduces to $\text{Var}_{S_4} \left(\frac{n_1 n_2}{m} \right)$ to the first term of a Taylor series, which according to Seber (1982:60) can be expressed as

$$N^2 \left(\frac{N}{n_1 n_2} + \frac{2N^2}{(n_1 n_2)^2} + \frac{6N^3}{(n_1 n_2)^3} \right). \quad (\text{A3})$$

The second term of Equation (A2) reduces to

$$E_{S_4} \left[\text{Var}_{S_3} \left(\frac{n_1 n_2}{\hat{m}} \mid S_4 \right) \right] = E_{S_4} \left[(n_1 n_2)^2 \text{Var}_{S_3} \left(\frac{1}{\hat{m}} \right) \right]$$

$$\begin{aligned}
E_{S_4} \left[\text{Var}_{S_3} \left(\frac{n_1 n_2}{\hat{m}} \mid S_4 \right) \right] &\doteq E_{S_4} \left[(n_1 n_2)^2 \frac{\text{Var}(\hat{m})}{m^4} \right] \text{ based on delta method} \\
&\doteq \text{Var}_{S_4} \left[\left(\frac{n_1 n_2}{m} \right)^2 \frac{\text{Var}(\hat{m})}{m^2} \right] \\
&\doteq \text{CV}(\hat{m})^2 \left[\text{Var}(\hat{N}) + E(\hat{N})^2 \right] \\
&\doteq \text{CV}(\hat{m})^2 \left[\text{Var}(\hat{N}) + N^2 \right]. \tag{A4}
\end{aligned}$$

The third term of Equation (A2) reduces to

$$\begin{aligned}
E_{S_4} \left[E_{S_3} \left[\text{Var}_{S_2} \left(\frac{n_1 \hat{n}_2}{\hat{m}} \mid S_3, S_4 \right) \right] \right] &= E_{S_4} \left[E_{S_3} \left(\frac{n_1}{\hat{m}} \right)^2 \text{Var}(\hat{n}_2) \right] \\
&\doteq E_{S_4} \left[\left(\frac{n_1}{m} \right)^2 \cdot \text{Var}(\hat{n}_2) \right] \text{ to 1st term of a Taylor series} \\
&\doteq \frac{n_2}{\left(\frac{n_1 n_2}{N} \right)^2} \text{Var}(\hat{n}_2) \text{ to 1st term of a Taylor series} \\
&\doteq N^2 \text{CV}(\hat{n}_2)^2. \tag{A5}
\end{aligned}$$

The fourth term of Equation (A2) reduces to

$$\begin{aligned}
E_{S_4} \left[E_{S_3} \left[E_{S_2} \left(\left(\frac{\hat{n}_2}{\hat{m}} \right)^2 \text{Var}(\hat{n}_1) \mid S_3, S_4 \right) \right] \right] &\doteq E_{S_4} \left[E_{S_3} \left(\left(\frac{\hat{n}_2}{\hat{m}} \right)^2 \text{Var}(\hat{n}_1) \mid S_4 \right) \right] \\
&\doteq E_{S_4} \left[\left(\frac{n_2}{m} \right)^2 \text{Var}(\hat{n}_1) \right] \text{ to 1st term of Taylor series} \\
&\doteq \frac{n_2}{\left(\frac{n_1 n_2}{N} \right)^2} \text{Var}(\hat{n}_1) \text{ to 1st term of a Taylor series} \\
&\doteq N^2 \text{CV}(\hat{n}_1)^2. \tag{A6}
\end{aligned}$$

Combining Equation (A3) to Equation (A5)

$$\begin{aligned}
\text{Var}(\hat{N}) &\doteq N^2 \left(\frac{N}{n_1 n_2} + \frac{2N^2}{(n_1 n_2)^2} + \frac{6N^3}{(n_1 n_2)^3} \right) + N^2 \text{CV}(\hat{n}_1)^2 + N^2 \text{CV}(\hat{n}_2)^2 \\
&\quad + (\text{Var}(\hat{N}) + N^2) \text{CV}(\hat{m})^2. \tag{A7}
\end{aligned}$$

Seber (1982:60) suggests the first term in Equation (A7) can be approximated by $N^2(1/m)$, in which case:

$$\text{Var}(\hat{N}) \doteq N^2 \left(\frac{1}{m} \right) + N^2 \cdot \text{CV}(\tilde{n}_1)^2 + N^2 \text{CV}(\hat{n}_2)^2 + (\text{Var}(\hat{N}) + N^2) \cdot \text{CV}(\hat{m})^2, \quad (\text{A8})$$

or yet more simply,

$$\text{Var}(\hat{N}) \doteq N^2 \left(\frac{1}{m} + \text{CV}(\hat{n}_1)^2 + \text{CV}(\hat{n}_2)^2 + \text{CV}(\hat{m})^2 \right). \quad (\text{A9})$$

Appendix B: Variance of \hat{m} , the Estimated Number of Mark Recaptures

The estimation of

$$\hat{m} = \frac{x}{\hat{p}_{\text{TR}}}, \quad (\text{B1})$$

where

x = actual tag recoveries observed,

\hat{p}_{TR} = probability of tag retention or operation.

In turn,

$$\hat{p}_{\text{TR}} = \frac{h_{12}}{h_1}, \quad (\text{B2})$$

where

h_{12} = recovered dual-tagged fish,

h_1 = recovered fish with spaghetti tag or dual tag.

$$\begin{aligned} \text{Var}(\hat{m}) &= \text{Var}\left(\frac{x}{\hat{p}_{\text{TR}}}\right) \\ &= \text{Var}_{S_2} \left[E_{S_1} \left(\frac{x}{\hat{p}_{\text{TR}}} \mid S_2 \right) \right] + E_{S_2} \left[\text{Var}_{S_1} \left(\frac{x}{\hat{p}_{\text{TR}}} \mid S_2 \right) \right], \end{aligned} \quad (\text{B3})$$

and where

S_1 = denotes estimation of \hat{p}_{TR} ,

S_2 = recovery of active tags.

$$\text{Var}(\hat{m}) = \text{Var}_{S_2} \left[\frac{x}{p_{\text{TR}}} \right] + E_{S_2} \left[x^2 \frac{\text{Var}(\hat{p}_{\text{TR}})}{p_{\text{TR}}^4} \right].$$

Assuming

$E\left(\frac{1}{\hat{p}_{\text{TR}}}\right) \doteq \frac{1}{p_{\text{TR}}}$, first-term Taylor series approximation, and

$\text{Var}\left(\frac{1}{\hat{p}_{\text{TR}}}\right) \doteq \text{Var}(\hat{p}_{\text{TR}})/p_{\text{TR}}^4$, Delta method,

$$\begin{aligned}\text{Var}(\hat{m}) &\doteq \frac{1}{p_{\text{TR}}^2} \text{Var}(x) + \frac{\text{Var}(\hat{p}_{\text{TR}})}{p_{\text{TR}}^4} E(x^2) \\ \text{Var}(\hat{m}) &\doteq \frac{1}{p_{\text{TR}}^2} \text{Var}(x) + \frac{\text{Var}(\hat{p}_{\text{TR}})}{p_{\text{TR}}^4} [\text{Var}(x) + E(x)^2].\end{aligned}$$

Note:

$$E(x) = mp_{\text{TR}},$$

$$\text{Var}(x) = mp_{\text{TR}}(1 - p_{\text{TR}}).$$

$$\therefore \text{Var}(\hat{m}) \doteq \frac{m(1 - p_{\text{TR}})}{p_{\text{TR}}} + \text{Var}(\hat{p}_{\text{TR}}) \left[\frac{mp_{\text{TR}}(1 - p_{\text{TR}}) + m^2 p_{\text{TR}}^2}{p_{\text{TR}}^4} \right]. \quad (\text{B4})$$

Note:

$$\text{Var}(\hat{p}_{\text{TR}}) = \frac{p_{\text{TR}}(1 - p_{\text{TR}})}{h_1}.$$

$$\therefore \text{Var}(\hat{m}) \doteq \frac{m(1 - p_{\text{TR}})}{p_{\text{TR}}} + \frac{m(1 - p_{\text{TR}})(mp_{\text{TR}} - p_{\text{TR}} + 1)}{p_{\text{TR}}^2 h_1} \quad (\text{B5})$$

estimated by

$$\widehat{\text{Var}}(\hat{m}) \doteq \frac{\hat{m}(1 - \hat{p}_{\text{TR}})}{\hat{p}_{\text{TR}}} + \frac{\hat{m}(1 - \hat{p}_{\text{TR}})(\hat{m}\hat{p}_{\text{TR}} - \hat{p}_{\text{TR}} + 1)}{\hat{p}_{\text{TR}}^2 h_1}. \quad (\text{B6})$$

Note a similar approach can be used to estimate the variance of \hat{n}_1 in 2015.

Appendix 3. Memo from AYK SSI Expert Panel to ADFG following preliminary review of revised run-reconstruction model following collaborative workshop between Expert Panel and ADFG to discuss original model limitations. Dated May 10, 2018. Please note that in this memo the panel referred to the 'original model' as the 'current model'.

MEMO

DATE: May 10, 2018

TO: Zachary Liller, Research Coordinator, Arctic-Yukon-Kuskokwim Region, Alaska
Department of Fish & Game, Division of Commercial Fisheries, Anchorage, Alaska

FROM: Expert Panel to evaluate Kuskokwim River Chinook salmon run reconstruction and stock-recruit models commissioned by the Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative (AYK-SSI).

Daniel Schindler, Professor, University of Washington, School of Aquatic and Fishery Sciences

Timothy Walsworth, Post-Doctoral Researcher, University of Washington, School of Aquatic and Fishery Sciences

Milo Adkison, Professor, College of Fisheries and Ocean Sciences, University of Alaska Fairbanks

Randall Peterman, Professor School of Resource and Environmental Management, Simon Fraser University

André Punt, Professor, University of Washington, School of Aquatic and Fishery Sciences

SUBJECT: Preliminary assessment of revised run reconstruction model for Chinook salmon in the Kuskokwim River

Introduction

Stocks of Chinook salmon returning to the Kuskokwim River are among the most abundant in Alaska but have shown downturns in the recent decade, resulting in closed commercial fisheries and hardship for subsistence fisheries in communities throughout the watershed. Stock assessments are particularly challenging in this large and remote river system because it is expensive and logistically difficult to detect and enumerate adult fish migrating from the ocean back to a complex network of spawning habitat distributed among the many tributaries of this river. A run reconstruction model is used by the Alaska Department of Fish & Game (ADF&G) to integrate among a variety of indices of abundance, including:

aerial surveys of spawning fish in headwater tributaries, counts of fish passing weirs on tributaries, and commercial catch rates in the lower river. Additionally, in some years, mark-recapture experiments are performed to estimate river-wide population abundance and provide a means for scaling from abundance indices to whole-system estimates in years where mark-recapture studies have not been done.

In response to concerns from a variety of stakeholders about the performance of the ADF&G run reconstruction model, the Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative (AYK-SSI) commissioned an independent panel of experts (hereafter Expert Panel), with considerable experience in salmon ecology and stock assessment, to review the structure and performance of the ADF&G's current published run reconstruction model (Bue et al 2012; hereafter 'current model'). The Expert Panel was assembled in 2016 and, combined with the work of a statistical analyst, initiated a collaborative review with the ADF&G to assess the performance of the current ADF&G run reconstruction model.

The Expert Panel used two approaches to assess the performance of the ADF&G run reconstruction model for Chinook salmon on the Kuskokwim River: (1) fitting the run reconstruction model to the observed data supplied by ADF&G, but with various modifications to that model's structure, and (2) fitting ADF&G's current run reconstruction model, including modified versions of it, to simulated data sets where the parameter values and run sizes are specified to simulate alternative plausible states of nature for the Kuskokwim River. A limitation of examining model performance on observed data is that the true state of the system is never known, and so there is no way to assess whether the model is actually capturing the true underlying dynamics in the system. Simulations allow for testing the model under various scenarios while being able to compare model fits to true values (Hilborn and Walters 1992).

The Expert Panel tested the current run reconstruction model in several ways to assess its sensitivity to the starting values for the parameters, to underlying assumptions about Chinook salmon population dynamics, and to the types and amounts of data used to estimate the model parameters. A thorough summary of these results will be available in a forthcoming Expert Panel Review expected to be completed in late May 2018. However, the primary conclusions of the Expert Panel were communicated at a collaborative workshop with ADF&G staff and their Kuskokwim River Interagency Chinook Salmon Run Reconstruction Model Development Team in March 2018, and a list of primary recommendations were made to improve model performance. In particular, the Expert Panel was concerned with:

- a) Lack of stability of the current run reconstruction model as demonstrated by its tendency to arrive at multiple solutions for the best values for the parameters of the model, depending on the starting values used in the model fitting process. Further investigation by the Expert Panel suggested that this instability derived from (1) an improperly specified harvest sub-model, and (2) over-parameterization of the escapement indices used to inform the model.
- b) Sensitivity of model estimates to inclusion of recent (2014-2017) mark-recapture data. The run reconstruction model produced substantially different estimates of historical run sizes when recent mark-recapture estimates were either used, or not, to anchor the run reconstruction effort.
- c) Error structure. The current model assumed a normal distribution for errors associated with the total run estimate derived from the mark-capture data and the Panel thought this would be better assumed to be log-normally distributed. The current model assumed that errors associated with

the individual escapement indices were distributed according to a negative binomial distribution, and each individual index site was assigned its own over-dispersion parameter. The Panel concluded that these errors should instead be assumed to be log-normally distributed and that the variances should be pooled by index type (i.e., one describing weirs and one describing aerial survey sites) to reduce the model complexity.

Following the Expert Panel's collaborative workshop in March 2018, ADF&G revised the run reconstruction model to account for several mutually agreed-upon revisions that the Panel suggested for improving model performance (Table 1).

Table 1. Comparison between current and revised model structures for ADF&G Kuskokwim River Chinook salmon run reconstruction model, as of May 1, 2018.

| Component | Current Model | Revised Model |
|---------------------------------------|--|--|
| Total Run Error Structure | Normal | Log-normal |
| Escapement Index Error Structure | Negative Binomial | Log-normal |
| Number of Escapement Error Parameters | One for each index site (20 total) | One for each type of index (2 total) |
| Harvest Component | Saturating relationship with effort: Catch \sim Run * (1-exp(-Effort*catchability)) | Linear relationship with effort: Catch \sim Effort * catchability * Run |

At the request of ADF&G, the Expert Panel performed a preliminary assessment of the performance of the revised run reconstruction model that was provided by ADF&G to the Panel on May 1, 2018. The purpose of this memo is to describe the results of this preliminary assessment. Given the short time frame, the Expert Panel was not able to perform an exhaustive assessment of the revised model but, instead, focused on a manageable number of critical concerns that emerged from the review of the current model as described above. For the purposes of this memo, we refer to the current model as the ‘current model’ and the revised model as the ‘revised model’. In reality, the core structure of these two models is fundamentally the same, but certain components have been revised in the new model provided on May 1, 2018.

Assessment of the revised model with historical observed data

Model stability

The revised model showed substantially improved stability compared to the current model as shown by less sensitivity to starting values for the initial run size (inset panels in Figure 1). While the current model settled on several local minima across the run reconstruction times-series (Figure 1 bottom panels), with and without the recent (2014-2017) mark-recapture data, the new model produced a single solution when all recent mark-recapture data were integrated into the run reconstruction (Figure 1, top right panel). The new model produced one renegade solution when the recent mark-recapture data were not used in the run reconstruction model (Figure 1, top left panel), but otherwise converged on a single solution.

Based on these preliminary analyses, it appears that model stability was substantially improved by the combination of simplifying the error structure by pooling many of the parameters and changing the harvest component of the model. While the revised model still showed some worrisome local minima when recent mark-recapture data were not included (Fig. 1 top, left panel), the revisions seem to have

distinctly improved model stability, particularly when recent (2014-2017) mark-recapture data are used in the run reconstruction. For future revisions to the model, the Expert Panel strongly recommends that ADF&G conduct simulation tests such as these to determine whether the run reconstruction model is sensitive to starting conditions. That procedure would examine model fits across a range of starting parameter values to ensure that a global minimum is found.

Influence of recent mark-recapture data

Mark-recapture estimates of river-wide abundance are needed to scale up from the miscellaneous escapement indices (i.e., weirs and aerial surveys of tributaries), which are assumed to quantify relative trends in abundance, to river-wide estimates of abundance. The Expert Panel noted that the run reconstruction estimates derived from using the current model were highly sensitive to the inclusion of recent (2014-2017) mark-recapture estimates of total river-wide abundance. The revised model remains sensitive to the inclusion of these data (Figure 2), though to a lesser degree than the current model. While the historical changes in abundance estimated from the current and revised models, with differing numbers of years of mark-recapture data, all generally followed the same coarse-scale changes through time, there were some notable discrepancies produced in certain years. In particular, the revised model generally tended to estimate lower total abundance of Chinook salmon between 2014-2017 than the current model did without using recent mark-recapture data for those years, but about the same as when the current model was fit using those data (Figure 2). Regardless, these differences in estimates were relatively small. The revised model also estimated the peak abundance observed in 1990s at more than 400,000 Chinook salmon while the current model estimated abundances almost 50,000 fish lower.

We further explored the sensitivity of the revised run reconstruction model to the inclusion of recent mark-recapture data by varying the number of years of mark-recapture data between 2014 and 2017 used in the run reconstruction. Given that there are no mark-recapture studies planned for 2018 and the following few years, this exercise is one way to assess how robust future estimates might be in years immediately following a series of mark-recapture estimates of river-wide abundance.

From 2010 – 2017, the revised model using all mark-recapture estimates during 2014-2017 estimated between a high of 133.3 thousand fish in 2017 to a low of 79.4 thousand fish in 2012 (Table 2a, right panel). When all four years of recent mark-recapture data were used in the run reconstruction, the deviations of the current model from the revised model estimates tended to be <5%, except for in 2014 when the current model estimated about 12% more fish in the river than was estimated by the revised model (Table 2, right panels).

By comparison, when no new mark-recapture data were used, the current model tended to overestimate the number of fish in the river from 2010-2017 compared to estimates produced by the revised model with all mark-recapture data. The estimates produced from the current model without new mark-recapture data tended to be <10% different from estimates with the revised model and all mark-recapture data. The one exception was 2014 when the current model estimated > 30% more fish than the revised model with all mark-recapture data. By comparison, the revised model without mark-recapture data produced estimates of total abundance that tended to be <5% different from estimates of the revised model fit with all of the mark-recapture data, except for in 2014 where the revised model without mark-recapture data estimated about 14% more fish than the revised model with all the mark-recapture data. The large error in 2014 appears to have been produced by abnormally high counts at two of the weir sites.

Assuming that run-size estimates from the revised model with all recent mark-recapture data are the closest to the true values, estimation accuracy of ADF&G's revised model decreased as fewer years of mark-recapture data were included in the run reconstruction (Table 3). However, these deviations tended to be small, and were typically <5% different from estimates generated by the revised model with all years of mark-recapture data (Table 3b). The one exception to this pattern was in the revised model's estimates of total run size for 2014, when produced without using any mark-recapture data, or when only the most recent (2015-2017) three years of data were used. These estimates were about 13% higher (>10,000 fish) than the estimates produced by the revised model based on all the recent (2014-2017) mark-recapture data. When mark-recapture data were used starting in 2014 (Table 3, three right-most columns), deviations from the situation where all years of mark-recapture data were used were negligible (<3%). Thus, the revised model remains sensitive to the inclusion of recent mark-recapture data, but less so than the current model. The model is particularly sensitive to exclusion of mark-recapture from years with unusual escapement patterns (which drive large estimation errors, e.g., 2014), but these years are more likely to be captured when mark-recapture studies are undertaken with increasing frequency. Further, the model appears to provide robust estimates of river-wide abundance in the years immediately following a mark-recapture experiment, although the analyses we have used to quantify this are very preliminary

Assessment of the revised model performance based on simulated data

We used a simulation model (documented in detail in the Expert Panel's upcoming final report) to generate data that would produce a reasonable approximation to the dynamics observed in Chinook salmon in the Kuskokwim River. The simulation model assumed that there was considerable population structure such that the aggregate dynamics were composed of the sum of the dynamics of 40 individual stocks, 20 of which were monitored for escapement. Covariation among stocks was assumed to be relatively weak, as demonstrated by the lack of synchrony among annual weir counts and among aerial surveys. The model also simulated 'productivity regimes', whereby the per capita productivity at low population sizes could increase by 500% (or decrease by 80%) roughly every 20 years. The model then 'sampled' the data at the intensity that has actually been performed in the Kuskokwim River over the last four decades (data become more sparse farther back in time; see Figure 4 x-axis).

Because we know what the 'real' abundances are in the model simulations, we can assess how well ADF&G's revised and current run reconstruction models perform in capturing these values under a variety of assumptions about the nature of the population dynamics and the intensity of sampling. In particular, we were interested in the influence of mark-recapture studies on model performance, and how the presence of regime shifts in population productivity affected model performance.

The revised model performed better than the current model in estimating the true abundance of Chinook salmon in simulated data (Figure 3); these improvements were particularly prominent in simulations where no new mark-recapture data were included in the run reconstructions. In the absence of regime dynamics and when no mark-recapture data were included, model performance (measured by the normalized root mean squared error, NRMSE) was substantially better for the revised model compared to the current model. However, with new mark-recapture data included, the difference in the NRMSE produced by the two models was negligible. In simulations with regime changes, the revised model performed about as well (as indicated by the NRMSE), regardless of whether new mark-recapture data

were included, and the frequency distributions of errors were only slightly wider in situations with regime shifts than without those shifts, regardless of whether new mark-recapture data were included in the run reconstructions (Figure 3).

Inspection of time-series of the relative errors produced by the current and the revised model through time reinforces the conclusion that the performance of the revised model still depends on inclusion of recent mark-recapture data in the run reconstructions, but less so than the current model (Figure 4). As expected, the magnitude of the errors of model predictions increases as you proceed backwards through time and the coverage of escapement sampling decreases. Inclusion of recent mark-recapture data tended to reduce errors in the most recent decade of the analysis, though the revised model had distinctly smaller errors than the current model during the last decade for simulations where new mark-recapture data were not included in the run reconstruction.

Summary

Revisions to the ADF&G run reconstruction model for Chinook salmon on the Kuskokwim River appear to have remedied several of the primary concerns of the AYK-SSI Expert Panel. In particular, the revised model is far more stable than the current model, though its stability still depends on the inclusion of recent mark-recapture data for scaling up from individual abundance indices to river-wide abundance estimates. The revised model also appears to provide more accurate run estimates than the current model, particularly for years when no mark-recapture data are available for scaling up to river-wide abundances. More analyses are required to further assess how robust the model is, particularly in situations where abundance indices from tributary weirs or aerial surveys are omitted from the Kuskokwim monitoring program.

References

Bue, B. G., K. L. Schaberg, Z. W. Liller, and D. B. Molyneaux. 2012. Estimates of the historic run and escapement for the Chinook salmon stock returning to the Kuskokwim River, 1976–2011. Alaska Department of Fish & Game, Fishery Data Series No. 12-49, Anchorage.

Hilborn, R., and C.J. Walters. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics, and Uncertainty. Chapman Hall. New York.

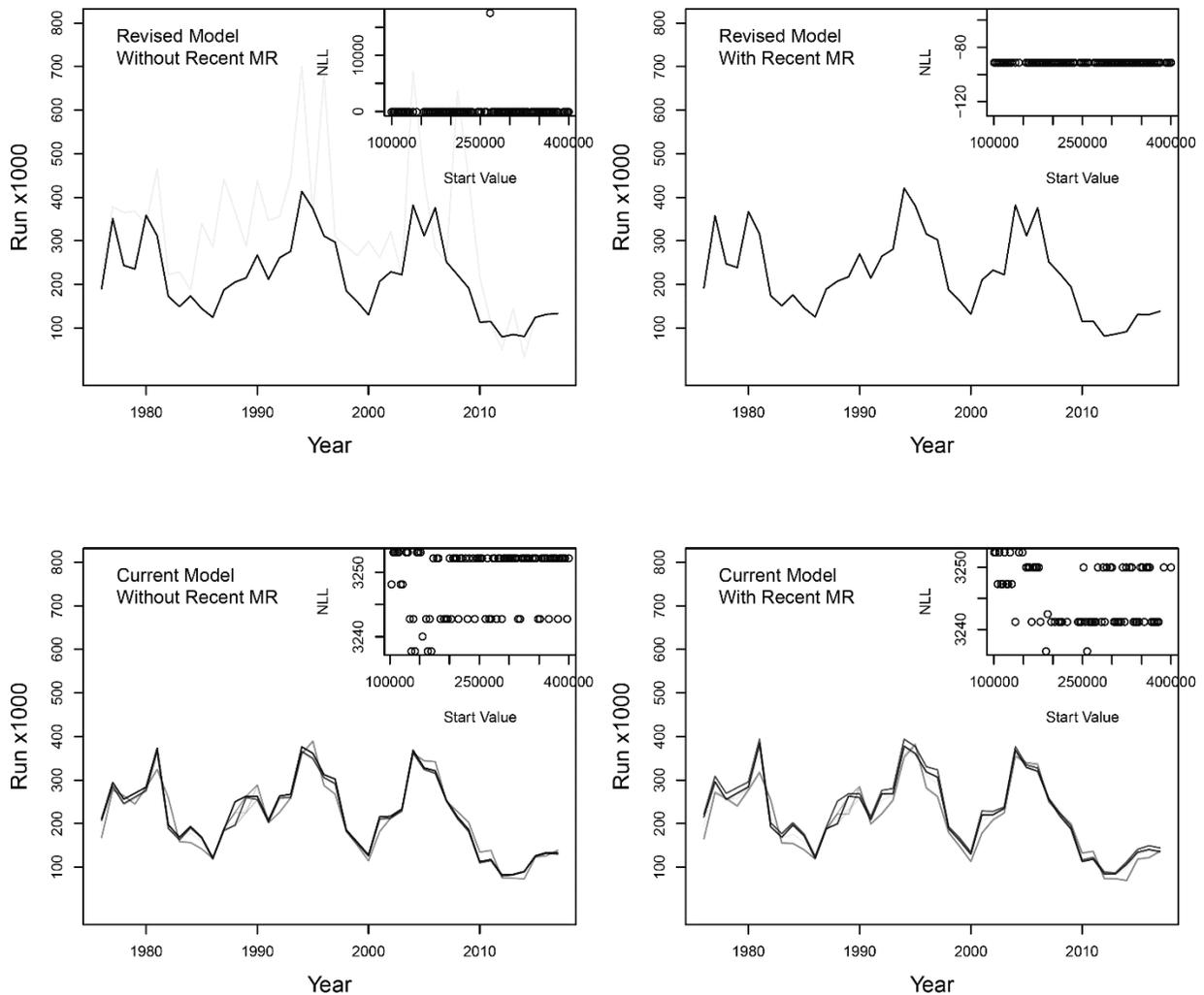


Figure 1. Run size estimates for Chinook salmon in the Kuskokwim River across a range of starting values from the revised run reconstruction model (top row) and current run reconstruction model (bottom row), and with different amounts of mark-recapture data available (no recent (2014-2017) estimates in left column, all recent estimates in right column). Semi-transparent grey lines represent individual model fits (out of 100 total). Black lines indicate stacked grey lines, representing repeated model convergence on the same values. Inset figures represent the negative log-likelihood values of model fits across the range of starting values of the run-size examined for the initial run size.

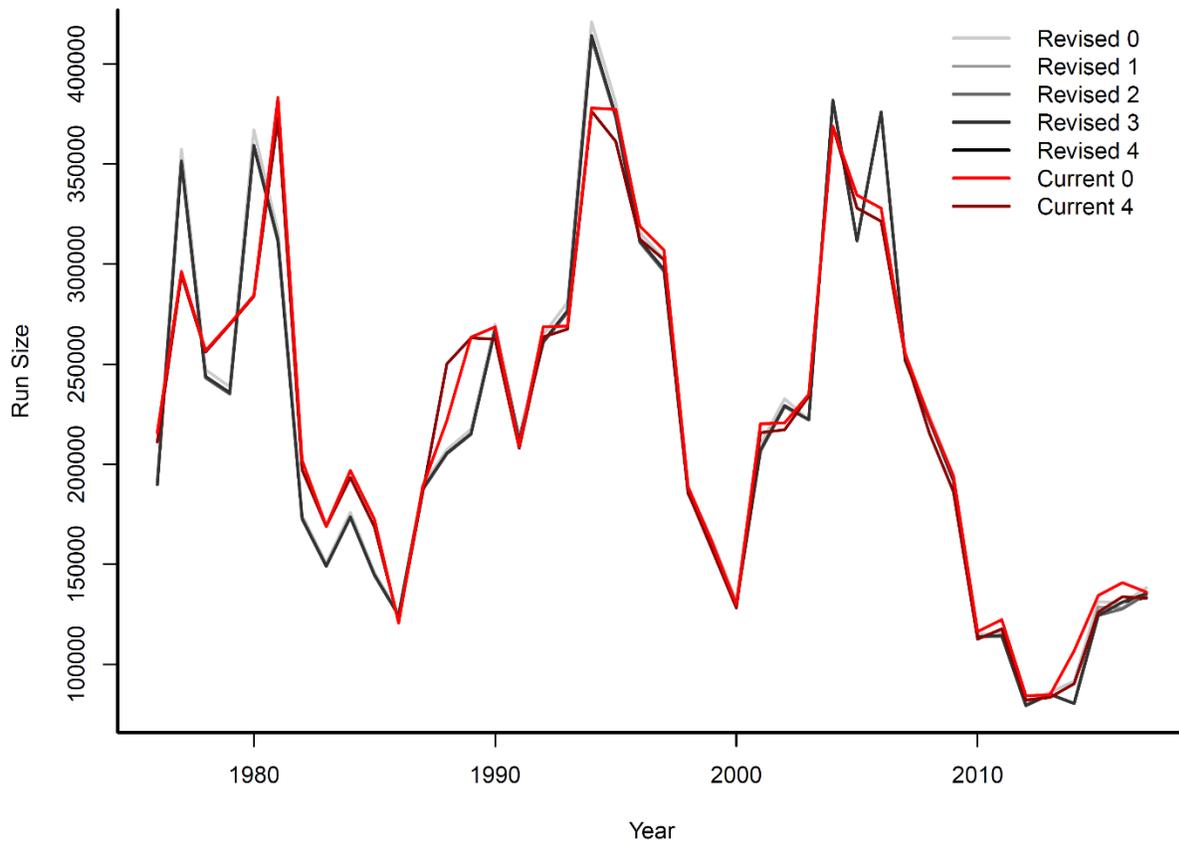


Figure 2. Point estimates of Kuskokwim River Chinook salmon run size using the current model (red and purple lines) and revised model (grey-scale lines) structures. The numbers in the legend following the model structure indicate the number of recent mark-recapture values used to fit the model (i.e., ‘Revised 0’ is the revised model fit without any mark-recapture data from 2014-2017. ‘Revised 4’ is the revised model fit with mark-recapture data for four years, 2014-2017. ‘Revised 1’ used only 2014 mark-recapture data, ‘Revised 2’ used only 2014 and 2015 mark-recapture data, and so on up through ‘Revised 4’.

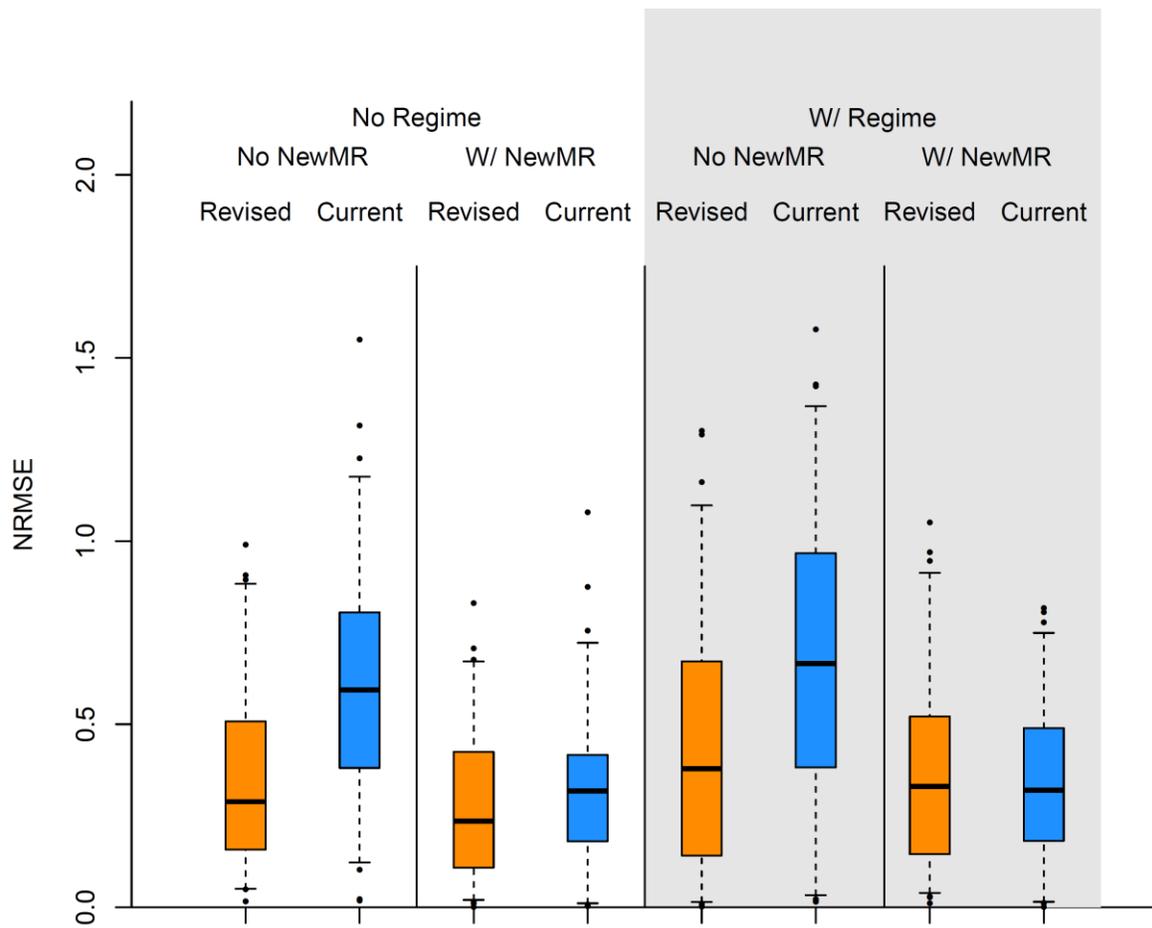


Figure 3. Boxplots of normalized root mean squared error (NRMSE) for Kuskokwim River Chinook salmon run reconstruction model fits to simulated data from an operating model under various biological scenarios and model structures. Box plots show the distribution from 100 simulations. The colors represent model estimates from the revised model structure (orange, left-most of each pair) and current model structure (blue, right-most). Column labels describe which model was used (Revised, Current), whether or not new (2014-2017) mark-recapture estimates were used to fit the models (No NewMR, W/ NewMR), and whether or not the underlying population dynamics were subject to regime shifts (also indicated by grey background).

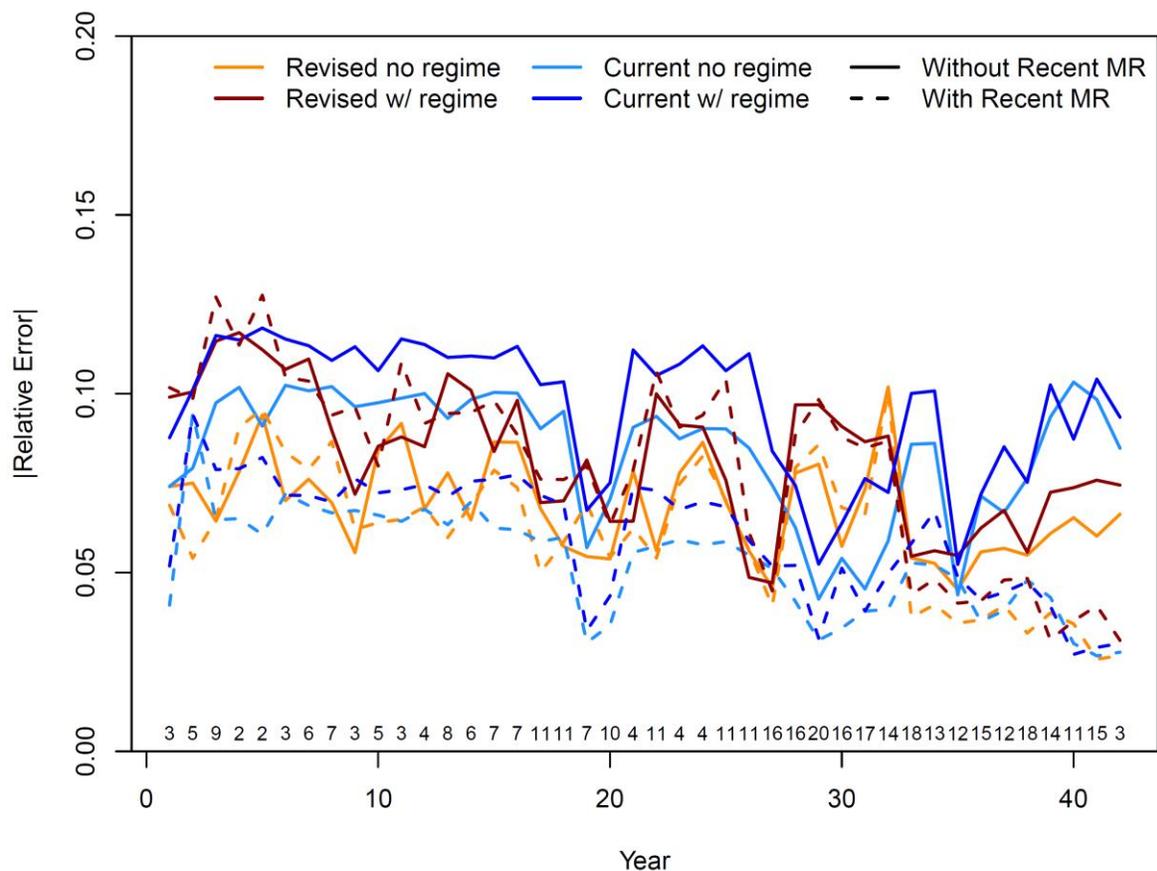


Figure 4. Median absolute values of relative error (expressed as proportional difference from the true value) through time in run reconstruction model estimates for 100 simulated time-series. Solid lines represent those in which the recent (2014-2017) mark-recapture estimates were not used in the run reconstruction model. Dashed lines represent scenarios in which the recent mark-recapture estimates were used in the run reconstruction model. Lines in orange shades represent results from the revised run reconstruction model, while blue shaded lines represent those from the current run reconstruction model. Darker shades of each color represent scenarios with population dynamics subject to regime shifts, while lighter shades represent scenarios without regime shifts. Numbers above x-axis indicate the number of escapement indices available each year, which are the same as in the real data set available for the Kuskokwim River.

Table 2. Comparisons of estimates of Kuskokwim River Chinook salmon abundance (run size in thousands of fish) from run reconstruction models using the revised and current model structures, and mark-recapture estimates of river-wide abundances. (a) Point estimates of Chinook salmon abundance from each of the two models when there are no recent mark-recapture estimates used and when there are all four recent mark-recapture estimates used. Grey boxes indicate years in which mark-recapture estimates are available. (b) Proportional differences between model estimates from part (a) compared to the revised model estimates when all recent mark-recapture estimates are used in the run reconstruction. Proportional differences were calculated as $[(\text{run size}_{\text{model } i} - \text{run size}_{\text{model } j}) / (\text{run size}_{\text{model } j})]$, where model j is the analogous ‘revised model’ fit with all (2014-2017) mark-recapture data.

| a) | No Recent Mark-Recapture | | All Recent Mark-Recapture | |
|------|--------------------------|---------------|---------------------------|---------------|
| | Revised Model | Current Model | Revised Model | Current Model |
| 2010 | 114.9 | 116.4 | 113.7 | 112.6 |
| 2011 | 115.7 | 122.3 | 114.3 | 117.7 |
| 2012 | 81.2 | 84.3 | 79.4 | 82.2 |
| 2013 | 86.0 | 84.8 | 85.0 | 83.5 |
| 2014 | 91.6 | 106.8 | 80.5 | 90.3 |
| 2015 | 131.3 | 134.4 | 124.4 | 126.1 |
| 2016 | 130.6 | 140.8 | 131.1 | 133.7 |
| 2017 | 138.3 | 136.1 | 133.3 | 133.1 |

| b) | No Recent Mark-Recapture | | All Recent Mark-Recapture | |
|------|--------------------------|---------------|---------------------------|---------------|
| | Revised Model | Current Model | Revised Model | Current Model |
| 2010 | 0.010 | 0.023 | 0.000 | -0.010 |
| 2011 | 0.012 | 0.071 | 0.000 | 0.030 |
| 2012 | 0.022 | 0.061 | 0.000 | 0.035 |
| 2013 | 0.011 | -0.003 | 0.000 | -0.018 |
| 2014 | 0.139 | 0.327 | 0.000 | 0.123 |
| 2015 | 0.055 | 0.080 | 0.000 | 0.014 |
| 2016 | -0.004 | 0.074 | 0.000 | 0.020 |
| 2017 | 0.037 | 0.021 | 0.000 | -0.001 |

Table 3. Comparisons of Kuskokwim River Chinook salmon run reconstruction estimates using the revised model structure and observed data, with different numbers of recent mark-recapture estimates available. (a) Point estimates of run size (thousands of fish) from the model fits with different numbers and arrangements of recent mark-recapture estimates used. Grey cells indicate years in which mark-recapture estimates were included in the run reconstruction. (b) Proportional differences (calculated as in Table 2) between all model estimates from (a) compared to the new model estimates when all recent mark-recapture estimates were used in the run reconstruction. Blue shading indicates underestimates; red shading indicates overestimates.

a)

| Year | Number of Recent Mark-Recapture Estimates Used | | | | | | | |
|------|--|-----------------|-------|-------|--------------------|-------------------|-------|-------|
| | No Estimates 0 | Later Estimates | | | All Estimates 4 | Earlier Estimates | | |
| | | 1 | 2 | 3 | | 3 | 2 | 1 |
| 2010 | 114.9 | 114.5 | 114.6 | 114.3 | 113.7 | 113.8 | 113.7 | 113.9 |
| 2011 | 115.7 | 115.2 | 115.3 | 115.0 | 114.3 | 114.4 | 114.2 | 114.4 |
| 2012 | 81.2 | 80.7 | 80.8 | 80.4 | 79.4 | 79.6 | 79.3 | 79.6 |
| 2013 | 86.0 | 85.7 | 85.7 | 85.5 | 85.0 | 85.1 | 85.0 | 85.1 |
| 2014 | 91.6 | 91.1 | 91.2 | 90.8 | 80.5 | 80.5 | 80.5 | 80.5 |
| 2015 | 131.3 | 130.4 | 130.6 | 124.4 | 124.4 | 124.4 | 124.4 | 128.6 |
| 2016 | 130.6 | 129.7 | 131.1 | 131.1 | 131.1 | 131.1 | 127.7 | 128.2 |
| 2017 | 138.3 | 133.3 | 133.3 | 133.3 | 133.3 | 135.5 | 135.0 | 135.5 |

b)

| Year | Number of Recent Mark-Recapture Estimates Used | | | | | | | |
|------|--|-----------------|-------|-------|--------------------|-------------------|--------|--------|
| | No Estimates 0 | Later Estimates | | | All Estimates 4 | Earlier Estimates | | |
| | | 1 | 2 | 3 | | 3 | 2 | 1 |
| 2010 | 0.010 | 0.007 | 0.008 | 0.005 | 0.000 | 0.001 | -0.001 | 0.001 |
| 2011 | 0.012 | 0.008 | 0.009 | 0.006 | 0.000 | 0.001 | -0.001 | 0.001 |
| 2012 | 0.022 | 0.015 | 0.017 | 0.012 | 0.000 | 0.002 | -0.002 | 0.002 |
| 2013 | 0.011 | 0.008 | 0.009 | 0.006 | 0.000 | 0.001 | -0.001 | 0.001 |
| 2014 | 0.139 | 0.132 | 0.134 | 0.128 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2015 | 0.055 | 0.048 | 0.050 | 0.000 | 0.000 | 0.000 | 0.000 | 0.033 |
| 2016 | -0.004 | -0.010 | 0.000 | 0.000 | 0.000 | 0.000 | -0.026 | -0.022 |
| 2017 | 0.037 | 0.000 | 0.000 | 0.000 | 0.000 | 0.016 | 0.012 | 0.017 |

Appendix 4. Computer code (R) for operating model used to generate simulated data for assessing performance of run-reconstruction models.

```
#=====
#
# Functions used in Operating Model for simulating
# population dynamics similar to those of the
# Kuskokwim Chinook salmon population, but accounting for
# sub-population dynamics
#
# Coded by: Timothy Walsworth, timothy.walsworth@usu.edu
#
#=====

#=====
# Load required libraries
#=====
library(gtools)
library(matrixcalc)
library(mvtnorm)
library(RColorBrewer)
library(Matrix)
library(R2admb)

#=====
# Function to generate the simulated time series data.
# - Simulates population dynamics of npop sub-populations
# - Each sub-population has its own production function
# - Harvest is distributed proportionally among
```

```
# sub-populations
# - Observations of escapement, harvest, and total run size
# are made with a specified amount of observation error
#
# Input parameters:
# - npop: number of sub-populations
# - nage: number of age classes
# - nwk: number of weeks simulated for harvest dynamics
# - nyr: number of years of data monitoring
# - oberr: log-normal observation error sd for aerial surveys
# - harvObsErr: log-normal sd of harvest observation error
# - procerr: recruitment error log-normal sd
# - rho: temporal autocorrelation in recruitment anomalies
# - corrtype: type of sub-population correlation structure
# - q: fishery catchability coefficient
# - mrstd: mark-recapture CV
# - tmrstd: total run mark recapture CV
# - agevar: (T/F) is there variation in age-composition between brood-years?
# - srtype: Shape of stock-recruit curve ("ricker" or "bevholt")
# - random.timing: (T/F) is run-timing variable among years?
# - lowpop: How many sub-populations in lower river? (Note: model currently
# does not do anything with this except when divying up weirs/surveys)
# - outnames: names for output matrix
# - decaymin: sets level of asynchrony among stocks (1= perfect synchrony, -1 -> mean cor = 0.29)
# - regime.length: inverse of annual probability of regime change
# - newmr: Include mark-recapture estimates from recent project?
# - aerial.bias: are aerial surveys increasingly biased at larger escapements?
# - varyEffort: randomly vary effort among years (to explore SR space better)
# - mryrs: Which years use mrstd for mark-recap estimates>
```

```

# - tmryrs: Which years use tmrstd for mark-recap estimates?
# - mrmn: 0 -> mark-recapture point estimate is the true run-size,
#     1 -> mar-recapture point estimate is drawn from distribution around true value
# - forecast: T/F: forecast annual harvest across range of harvest rates?
# - regstr: multiplier for difference in productivity between low and high productivity regimes
# - avar: T/F, is there variation in population productivity among populations?

#=====
=====

generate.time.series<-function(npop,nage,nwk,nyr,oberr,
                               harvObsErr,procerr,rho,corrtype="none",q,
                               mrstd,tmrstd,agevar=F,qvar=F,srtype="ricker",random.timing=F,lowpop=10,
                               outnames,decaymn=0,regime.length,newmr=F,aerial.bias=F,varyEffort=F,
                               mryrs,tmryrs,mrmn,forecast=F,regstr=2,avar=F)
{
  # Read in Kusko data file in order to determine which weeks have harvest and fishing effort
  data_file2 <- 'Kusko_RR_Input_Data_April2018.csv'
  kusko.data2 <- read.csv(data_file2,header=T, na.string="")

  #=====

  # Generate input files in order to determine which estimates are available
  # - Start Code from Hamazaki's R code for Kusko RR model

  #=====

  # Extract testfish data
  testf<-kusko.data2[substr(names(kusko.data2),1,3)=='rpw']
  # combine week 8, 9 and 10 and drop
  testf[,8] <- testf[,8]#+testf[,9]+testf[,10]
  #testf <- testf[,-(9:10)]

  # Replace NA to mean proportion for each week
  for (i in 1:dim(testf)[2]) {
    testf[is.na(testf[i]),i] <- colMeans(testf,na.rm=T)[i]
  }
}

```

```
}
```

```
#####
```

```
####
```

```
# 2.3 Rearrange fishing effort and harvest data catch 0 to NA #
```

```
#####
```

```
####
```

```
# Extract weekly commercial effort data
```

```
ceff2 <-kusko.data2[substr(names(kusko.data2),1,3)=='cew']
```

```
# combine week 8, 9 and drop
```

```
# ceff2[,6] <- ceff2[,6]+ceff2[,7]
```

```
ceff2 <- ceff2[,-7]
```

```
# Extract weekly commercial catch data
```

```
ccat2 <-kusko.data2[substr(names(kusko.data2),1,3)=='chw']
```

```
# combine week 8, 9 and drop
```

```
# ccat2[,6] <- ccat2[,6]+ccat2[,7]
```

```
ccat2 <- ccat2[,-7]
```

```
# Extract weekly commercial est data
```

```
creg2 <-kusko.data2[substr(names(kusko.data2),1,3)=='cfw']
```

```
# combine week 8, 9 and drop
```

```
creg2[,6] <- pmax(creg2[,6],creg2[,7])
```

```
creg2 <- creg2[,-7]
```

```
creg2[creg2==4]<-1
```

```
#=====
```

```
# End Code from Hamazaki's R code for Kusko RR model
```

```
#=====
```

```

# Generate Stock-recruit parameters for sub-stocks
if(srtype=="ricker"){

  alphas.hi<-3*regstr # High productivity regimes
  if(avar) alphas.hi<-rnorm(npop,(3*regstr),(.45*regstr)) # Set population specific high productivity
  alphas.lo<-3 # Low productivity regimes
  if(avar) alphas.lo<-rnorm(npop,3,.45) # Set population specific low productivity
  maxr<-rnorm(npop,25000,6000) # Maximum recruitment
  aset<-sample(c(1,2),1) # Which regime to start in?
  if(regime.length==0) aset<-1 # If no regime shifts, maintain same regime throughout simulation
  alphas<-cbind(alphas.hi,alphas.lo)[,aset] # Select alpha parameters for regime
  betas<--(alphas/maxr)*exp(-1) # Set betas based on max recruitment and productivity
  srparmsmat<-cbind(alphas.hi,alphas.lo,betas) # Matrix of SR parameters by population
}

if(srtype=="bevholt"){

  alphas<-rnorm(npop,25000,6000) # Max recruitment
  betaslo<-alphas/3 # Set low productivity regime beta
  if(avar) betaslo<-alphas/rnorm(npop,3,.45) # Set population specific low productivity
  betashi<-alphas/(3*regstr) # Set high productivity regime beta
  if(avar) betashi<-alphas/rnorm(npop,(3*regstr),(.45*regstr)) # Set population specific high
productivity
  aset<-sample(c(1,2),1) # Which regime?
  if(regime.length==0) aset<-1 # if no regime shifts, maintain same regime throughout
  betasmat<-cbind(betashi,betaslo) # combine high and low productivity parameters into matrix
  betas<-betasmat[,aset] # Select beta for current regime
  srparmsmat<-cbind(alphas,betashi,betaslo) # Matrix of SR parmameters by populations
}

```

```

# Generate productivity anomalies, declaring covariance structure among stocks

anomsout<-
generate.anoms(npop=npop,synchrony=corrtype,nyr=nyr+100,procerr=procerr,decaymin=decaymn) #
generate anomalies

anoms<-anomsout[[1]] # productivity anomalies (rows = years, columns = stock)

spatialcorr<-anomsout[[2]] # covariance structure among stocks

# Generate Starting Data

# Matrix with returns, escapement, harvest, recruits, recruits by age

popmat<-array(NA,dim=c(nyr,4+nage+1,npop)) # Dim 1: Year, Dim 2: Data columns, Dim 3: Population
ID

popmat[1:8,2,]<-exp(rnorm(8*npop,9,.5)) # Fill starting escapements from random normal distribution
# Columns of popmat for each population<-c("Ret","Esc","Harv","Rec","R5","R6","R7","R8","eps")
dimnames(popmat)<-list(c(seq(1,nyr)),
                      c("Ret","Esc","Harv","Rec","R5","R6","R7","R8","eps"),
                      c(seq(1:40)))

#selec<-c(.2,.8,1,1) # Age-specific selectivity of harvest
selec<-c(1,1,1,1) # Even selectivity across all return ages

if(!qvar){
  qmean<-q
  q<-rep(q,nyr-8) # Catchability -> Constant through time
}

#q<-rnorm(nyr,q,q*0.1) # Time-varying catchability

if(qvar) {
  qmean<-q
  q<-runif(nyr-8,.75*qmean,1.25*qmean)
}

```

```

}
TotBoats<-.5/qmean # Set effort level
if(varyEffort){
  TotBoats2<-as.matrix(runif(nyr-8,.5,3)*TotBoats,nrow=nyr,ncol=1) # Vary effort across years
}

harvdy<-array(NA,dim=c(nyr,8,npop,nwk)) # Store harvest data (Dim 1: Year, Dim 2: Harvest Metric,
Dim 3: Population, Dim 4: Week)

for(i in 1:8)
{

# Are regime shifts being simulated?
if(regime.length!=0){
  if(sample(c(rep(0,regime.length-1),1),1)) # Does a regime shift occur this year?
  {
    # Switch productivity parameters if regime shift occurs
    if(srtype=="ricker"){
      if(aset==1) alphas<-alphas.lo
      else alphas<-alphas.hi
      aset<-abs(aset-3) # Change indicator of regime from 1 to 2, or 2 to 1
    }
    if(srtype=="bevholt"){
      if(aset==1) betas<-betas.lo
      else betas<-betas.hi
      aset<-abs(aset-3) # Change indicator of regime from 1 to 2, or 2 to 1
    }
  }
}

```

```

}
}

# Generate and record Recruits for each spawning population
for(j in 1:npop)
{

  sr<-rick(spawn=popmat[i,2,j],sig=procerr,a=alphas[j],b=betas[j],rho=rho,
           last.eps=ifelse(i==1,0,popmat[i-1,9,j]),anom=anoms[i,j],srtype=srtype) # generate recruits for
population j

  popmat[i,4,j]<-sr[1] # Store recruits
  popmat[i,9,j]<-sr[2] # Store recruitment anomaly

  for(k in 1:nage)
  {

    popmat[i+k+4,k+4,j]<-age.comp(popmat[i,4,j],npop=1)[k] # Generate age composition for
population k

    if(!lagevar) popmat[i+k+4,k+4,j]<-(c(.19,.38,.390,.04)*popmat[i,4,j])[k] # Generate age composition
for population k without variation among years

  }

}

}

ret.time<-matrix(NA,nrow=nyr-8,ncol=10) # Storage for return timing
effort.time<-matrix(NA,nrow=nyr-8,ncol=7) # Storage for effort timing
if(!random.timing){ # is return timing variable across years

  ret.time<-matrix(returntime(npop=1),nrow=nyr-8,ncol=10,byrow=T) # Generate and store return
timing

```

```
effort.time<-matrix(effort(npop=1)*TotBoats,nrow=nyr-8,ncol=7,byrow=T) # Generate and store
effort timing
```

```
if(varyEffort) effort.time<-t(apply(TotBoats2,MARGIN=1,FUN=function(x) rep(x,7))) # IF effort varies
widely across years, use that year's effort across all weeks
```

```
effort.time[(nyr-58):(nyr-18),][is.na(creg2) | creg2==0 |creg2==4]<- 0 # Only exert fishing effort during
periods when real fishing effort was exerted
```

```
}
```

```
for(i in 9:(nyr-10))
```

```
{
```

```
# Are regime shifts being simulated?
```

```
if(regime.length!=0){
```

```
if(sample(c(rep(0,regime.length-1),1),1)) # Is there a regime shift this year?
```

```
{
```

```
# Switch productivity parameters if there is a regime shift
```

```
if(srtype=="ricker"){
```

```
if(aset==1) alphas<-alphas.lo
```

```
else alphas<-alphas.hi
```

```
aset<-abs(aset-3)
```

```
}
```

```
if(srtype=="bevholt"){
```

```
if(aset==1) betas<-betas.lo
```

```
else betas<-betas.hi
```

```
aset<-abs(aset-3)
```

```
}
```

```
}
```

```
}
```

```

# Across all sub-populations
for(j in 1:npop)
{
  popmat[i,1,j]<-sum(popmat[i,5:8,j]) # Add up returns across ages

  for(z in 1:nwk)
  {
    for(ageclass in 1:nage)
    {

      harvdy[i,ageclass,j,z]<-(ret.time[i-8,z+2]*popmat[i,4+ageclass,j])*(1-exp(-1*effort.time[i-
8,z]*q[i]*selec[ageclass])) # Calculate harvest by age and week

    }

  }

  popmat[i,3,j]<-sum(harvdy[i,1:4,j,]) # Harvest total in year i, population j
  popmat[i,2,j]<-popmat[i,1,j]-popmat[i,3,j] # Escapement total in year i, population j

  sr<-rick(spawn=popmat[i,2,j],sig=procerr,a=alphas[j],b=betas[j],
          rho=rho,last.eps=popmat[i-1,9,j],anom=anoms[i,j],srtype=srtype) # Calculate recruitment
dynamics
  popmat[i,4,j]<-sr[1] # Recruits from brood year i, population j
  popmat[i,9,j]<-sr[2] # epsilon for AR1 process, year i, population j
  agecomp<-age.comp(popmat[i,4,j],npop=1) # generate age composition for the current brood year
  for(k in 1:nage)

```

```

{
  popmat[i+k+4,k+4,j]<-agecomp[k] # distribute recruits from brood year i among age-classes in
  respective return years
  if(!agevar) popmat[i+k+4,k+4,j]<-(c(.19,.38,.390,.04)*popmat[i,4,j])[k] # Distribute recruits among
  age-classes when age composition is constant
}

}

}

#print("B")

obsmat<-array(NA,dim=c(nyr,12,npop)) # Storage for observations of the system

# obsmat columns: 1. Obsreturns, 2. ObsEscape, 3. ObsHarv, 4. ObsRecruits, 5:8. Obs Escape by age,
9:12. Obs Harv by Age

# Store harvest data (Dim 1: Year, Dim 2: Harvest Metric, Dim 3: Population, Dim 4: Week)

# Select sites for weirs and aerial indices (selects from lower and upper system, distributing weirs and
aerials as in Kusko)

monitored<-
c(sample(c(rep(0,5),1,1,2,2,2),lowpop,replace=F),sample(c(rep(0,15),rep(1,4),rep(2,11)),npop-
lowpop,replace=F))

aerial<-as.integer(monitored==2) # Which sites are aerial indices?

weir<-as.integer(monitored==1) # Which sites have weirs?

nweir<-sum(monitored==1) # How many populations have at least one weir index?

weirs<-which(monitored==1) # How many populations have at least one aerial index?

nair<-sum(monitored==2) # Number of weir index sites

air<-which(monitored==2) # number of aerial index sites

harvbyage<-array(NA,dim=c(nyr,nage,npop)) # Storage for harvest by age and population

```

```

escbyage<-array(NA,dim=c(nyr,nage,npop)) # Storage for escapement by age and populations

for(z in 9:(nyr-10))
{

for(i in 1:npop)
{

for(j in 1:nage)
{
escbyage[z,j,i]<-popmat[z,j+4,i]-(sum(harvdy[z,j,i])) # Store escapement by age
harvbyage[z,j,i]<-(sum(harvdy[z,j,i])) # Store harvest by age
}

if(i %in% weirs) {
obsmat[z,2,i]<-observe.weir(escape=popmat[z,2,i],obserr.weir=oberr*.2,weir=1,npop=1) #
Generate escapement "observations" at weir locations

}

# Generate escapement "observations" at aerial survey locations
if(i %in% air) obsmat[z,2,i]<-
observe.aerial(escape=popmat[z,2,i],obserr.aerial=oberr,aerial=1,npop=1,aerialbias=aerial.bias)

obsmat[z,3,i]<-popmat[z,3,i]*exp(rnorm(1,-1*(harvObsErr^2/2),harvObsErr)) # Store observed
harvest

obsmat[z,1,i]<-obsmat[z,2,i]+obsmat[z,3,i] # Store observed harvest + obsrved escapement

```

```

}

}

modintab<-matrix(NA,nrow=nyr-8,ncol=75) # Create storage for output matrix

# Set column headers for output matrix
if(length(outnames)==58)
{
  newcolnames<-c(outnames[1:3],"NoUSE",outnames[c(4:7)],rep("NOUSE",4),outnames[c(8:58)],
    "Age4Harvest","Age5Harvest","Age6Harvest","Age7Harvest","TRUE Run",
    "Age4Esc","Age5Esc","Age6Esc","Age7Esc","Returns","KuskoEsc","Effort")
  colnames(modintab)<-newcolnames

}

else colnames(modintab)<-
c(outnames,"Age4Harvest","Age5Harvest","Age6Harvest","Age7Harvest","TRUE Run",
  "Age4Esc","Age5Esc","Age6Esc","Age7Esc","Returns","KuskoEsc","Effort")

for(i in 1:(nyr-8))
{
  modintab[i,73]<-sum(popmat[i+8,4,]) # Recruits
  modintab[i,74]<-sum(popmat[i+8,2,]) # Escapement
  modintab[i,75]<-TotBoats2[9:nyr] # Effort
  modintab[i,68]<-sum(popmat[i+8,1,]) # Returns
  modintab[i,1]<-i # Year number

  modintab[i,2]<-sum(popmat[i+8,3,])* .1 # Harvest.up

```

```

modintab[i,3]<-sum(popmat[i+8,3,])*.9 # Harvest.low
modintab[i,4]<-NA # H. Sub.I
modintab[i,5]<-NA # H.Sports
modintab[i,6]<-sum(popmat[i+8,3,])*.00 # H.Test

for(j in 1:nage)
{
  modintab[i,(63+j)]<-sum(harvbyage[i+8,j,])/sum(harvbyage[i+8,,]) # Proportional harvest by age
}
for(j in 1:nage)
{
  modintab[i,(68+j)]<-sum(escbyage[i+8,j,])/sum(escbyage[i+8,,]) # Proportional Escapement by Age
}

if(i %in% tmryrs){
  lcv<-log(tmrstd^2+1) # log cv
  modintab[i,8]<- tmrstd*sum(popmat[i+8,1,]) # tmr.sd
  modintab[i,7]<- (mrmn*sum(popmat[i+8,1,])*exp(rnorm(1,-lcv^2/2,sqrt(lcv))))+((1-
mrmn)*sum(popmat[i+8,1,]))) # tmr
  modintab[i,9]<- NA # Sonar
  modintab[i,10]<- NA # Sonar.sd
}

if(i %in% mryrs)
{
  firstup<-lowpop+1 # Not used currently; would be incorporated if upper and lower
harvest/escapement dynamics are incorporated
  lcv<-log(mrstd^2+1) # log CV
  modintab[i,8]<-mrstd*sum(popmat[i+8,1,]) #changed from saving in column 12; removed firstup:40
from third index

```

```
modintab[i,7]<-(mrmn*sum(popmat[i+8,1,])*exp(rnorm(1,(-1*(lcv)^2)/2,sqrt(lcv))))+((1-  
mrmn)*sum(popmat[i+8,1,])) # All years actual run size; changed from saving in column 11
```

```
}
```

```
# Columns 13-18 are weir indices of escapement
```

```
for(j in 1:nweir){
```

```
  pop<-weirs[j] # which subpopulation to index?
```

```
  modintab[i,12+j]<-obsmat[i+8,2,pop] # Weir escapement
```

```
}
```

```
# Columns 19-32 are aerial indices of escapement
```

```
for(j in 1:nair){
```

```
  pop<-air[j] # which sub-population to index?
```

```
  modintab[i,18+j]<-obsmat[i+8,2,pop] # Aerial escapement
```

```
}
```

```
# Columns 33-42 are proportional return times for each year
```

```
modintab[i,33:42]<-ret.time[i,] # Proportion of run returning in week x
```

```
# columns 41,44,47,50,53,56,59 are Fishing harvest in week x
```

```
for(j in 1:nwk)
```

```
{
```

```
  modintab[i,(40+3*j)]<-sum(harvdy[i+8,1:4,,j]) # Harvest by week across all ages and populations and  
fisheries
```

```
}
```

```
# Columns 42,45,48,51,54,57,60 are effort in week x
```

```
for(j in 1:nwk)
```

```

{
  modintab[i,(41+3*j)]<-effort.time[i,j] # effort by week
}

# columns 43,46,49,52,55,58,61 are Fishing regulation in week x
for(j in 1:nwk)
{
  modintab[i,(42+3*j)]<-1 #Fishing regulation by week (only one regulation simulated)
}

}

substock.runs<-popmat[,1] # Substock run dynamics
substock.esc<-popmat[,2] # Sub-stock escapement dynamics

#=====
# Forecast 50 years across harvest rates
if(forecast)
{

  TotBoats3<- matrix(rep((-1*log(1-seq(.01,.99,by=.01))/q[1]),110),nrow=110,ncol=99,byrow=T) #Matrix
of effort for different harvest scenarios

  popmat.f<-array(NA,dim=c(118,4+nage+1,npop,99)) # 4d array for storing population forecast output
  harvdyf<-array(NA,dim=c(110,8,npop,nwk,99)) # 4d matrix for storing harvest dynamics forecast
output
  harvbyagef<-array(NA,dim=c(110,nage,npop,99)) # Storage for harvest by age and population
  escbyagef<-array(NA,dim=c(110,nage,npop,99)) # Storage for escapement by age and populations
  obsmatf<-array(NA,dim=c(110,12,npop,99)) # Storage for observations of the system
  harvbyagef<-array(NA,dim=c(110,nage,npop,99)) # Storage for harvest by age and population
  escbyagef<-array(NA,dim=c(110,nage,npop,99)) # Storage for escapement by age and populations

```

```

substock.runsf<-array(NA,dim=c(118,npop,99)) # storage for forecast substock run dynamics
substock.escf<-array(NA,dim=c(118,npop,99)) # storage for forecast substock escapement dynamics
ret.timef<-matrix(NA,nrow=110,ncol=10) # Storage for return timing
effort.timef<-TotBoats3 # Storage for effort timing
if(!random.timing){ # is return timing variable across years
  ret.timef<-matrix(returntime(npop=1),nrow=110,ncol=10,byrow=T) # Generate and store return
timing
}

for(vv in 1:99)
{
  popmat.f[1:16,,vv]<-popmat[(nyr-17):(nyr-2),,] # Set initial conditions for all forecast simulations
}

modintabf<-array(NA,dim=c(110-8,75,99)) # Create storage for output matrix

# Set column headers for output matrix
if(length(outnames)==58)
{
  newcolnames<-c(outnames[1:3],"NoUSE",outnames[c(4:7)],rep("NOUSE",4),outnames[c(8:58)],
    "Age4Harvest","Age5Harvest","Age6Harvest","Age7Harvest","TRUE Run",
    "Age4Esc","Age5Esc","Age6Esc","Age7Esc","Returns","KuskoEsc","Effort")
  colnames(modintabf)<-newcolnames
}

else colnames(modintabf)<-
c(outnames,"Age4Harvest","Age5Harvest","Age6Harvest","Age7Harvest","TRUE Run",
  "Age4Esc","Age5Esc","Age6Esc","Age7Esc","Returns","KuskoEsc","Effort")

```

```
if(regime.length !=0) regimetimes<-sample(c(rep(0,regime.length-1),1),110,replace=T) # Generate
vector of regime switch points
```

```
for(v in 1:99)
{
  for(t in 9:110)
  {
    set.seed(1)
    # Are regime shifts being simulated?
    if(regime.length!=0){
      #if(sample(c(rep(0,regime.length-1),1),1)) # Is there a regime shift this year?
      if(regimetimes[t])
      {
        # Switch productivity parameters if there is a regime shift
        if(srtype=="ricker"){
          if(aset==1) alphas<-alphas.lo
          else alphas<-alphas.hi
          aset<-abs(aset-3)
        }
        if(srtype=="bevholt"){
          if(aset==1) betas<-betas.lo
          else betas<-betas.hi
          aset<-abs(aset-3)
        }
      }
    }
  }
}
```

```
# Across all sub-populations
```

```

for(j in 1:npop)
{
  popmat.f[t,1,j,v]<-sum(popmat.f[t,5:8,j,v]) # Add up returns across ages

  for(z in 1:nwk)
  {
    for(ageclass in 1:nage)
    {

      harvdyf[t,ageclass,j,z,v]<-(ret.timef[t-8,z+2]*popmat.f[t,4+ageclass,j,v])*(1-exp(-
1*effort.timef[t-8,v]*q[t]*selec[ageclass])) # Calculate harvest by age and week

    }

  }

  popmat.f[t,3,j,v]<-sum(harvdyf[t,1:4,j,,v]) # Harvest total in year i, population j
  popmat.f[t,2,j,v]<-popmat.f[t,1,j,v]-popmat.f[t,3,j,v] # Escapement total in year i, population j

  sr<-rick(spawn=popmat.f[t,2,j,v],sig=procerr,a=alphas[j],b=betas[j],
          rho=rho,last.eps=popmat.f[t-1,9,j,v],anom=anoms[t,j],srtype=srtype) # Calculate recruitment
dynamics
  popmat.f[t,4,j,v]<-sr[1] # Recruits from brood year i, population j
  popmat.f[t,9,j,v]<-sr[2] # epsilon for AR1 process, year i, population j
  agecomp<-age.comp(popmat.f[t,4,j,v],npop=1) # generate age composition for the current brood
year
  for(k in 1:nage)
  {
    popmat.f[t+k+4,k+4,j,v]<-agecomp[k] # distribute recruits from brood year i among age-classes
in respective return years

```

```
    if(!agevar) popmat.f[t+k+4,k+4,j,v]<-(c(.19,.38,.390,.04)*popmat.f[t,4,j,v])[k] # Distribute recruits
among age-classes when age composition is constant
```

```
  }
```

```
}
```

```
}
```

```
for(z in 9:110)
```

```
{
```

```
  for(i in 1:npop)
```

```
  {
```

```
    for(j in 1:nage)
```

```
    {
```

```
      escbyagef[z,j,i,v]<-popmat.f[z,j+4,i,v]-(sum(harvdyf[z,j,i,,v])) # Store escapement by age
```

```
      harvbyagef[z,j,i,v]<-(sum(harvdyf[z,j,i,,v])) # Store harvest by age
```

```
    }
```

```
    if(i %in% weirs) {
```

```
      obsmatf[z,2,i,v]<-observe.weir(escape=popmat.f[z,2,i,v],obserr.weir=oberr*.2,weir=1,npop=1) #
```

```
Generate escapement "observations" at weir locations
```

```
    }
```

```

# Generate escapement "observations" at aerial survey locations

if(i %in% air) obsmatf[z,2,i,v]<-
observe.aerial(escape=popmat.f[z,2,i,v],obserr.aerial=oberr,aerial=1,npop=1,aerialbias=aerial.bias)

    obsmatf[z,3,i,v]<-popmat.f[z,3,i,v]*exp(rnorm(1,-1*(harvObsErr^2/2),harvObsErr)) # Store
observed harvest

    obsmatf[z,1,i,v]<-obsmatf[z,2,i,v]+obsmatf[z,3,i,v] # Store observed harvest + obsrved
escapement

}

}

for(i in 1:(110-8))
{
    modintabf[i,73,v]<-sum(popmat.f[i+8,4,,v]) # Recruits
    modintabf[i,74,v]<-sum(popmat.f[i+8,2,,v]) # Escapement
    modintabf[,75,v]<-TotBoats3[9:110] # Effort
    modintabf[i,68,v]<-sum(popmat.f[i+8,1,,v]) # Returns
    modintabf[i,1,v]<-i # Year number

    modintabf[i,2,v]<-sum(popmat.f[i+8,3,,v])*0.1 # Harvest.up
    modintabf[i,3,v]<-sum(popmat.f[i+8,3,,v])*0.9 # Harvest.low
    modintabf[i,4,v]<-NA # H. Sub.I
    modintabf[i,5,v]<-NA # H.Sports
    modintabf[i,6,v]<-sum(popmat.f[i+8,3,,v])*0.00 # H.Test

    for(j in 1:nage)

```

```

    {
      modintabf[i,(63+j),v]<-sum(harvbyagef[i+8,j,,v])/sum(harvbyagef[i+8,,,v]) # Proportional harvest
by age
    }
    for(j in 1:nage)
    {
      modintabf[i,(68+j),v]<-sum(escbyagef[i+8,j,,v])/sum(escbyagef[i+8,,,v]) # Proportional
Escapement by Age
    }

    if(i %in% tmryrs){
      lcv<-log(tmrstd^2+1)
      modintabf[i,8,v]<- tmrstd*sum(popmat.f[i+8,1,,v]) # tmr.sd
      modintabf[i,7,v]<- (mrmn*sum(popmat.f[i+8,1,,v])*exp(rnorm(1,-lcv^2/2,sqrt(lcv))))+(((1-
mrmn)*sum(popmat.f[i+8,1,,v]))) # tmr
      modintabf[i,9,v]<- NA # Sonar
      modintabf[i,10,v]<- NA # Sonar.sd
    }

    if(i %in% mryrs)
    {
      firstup<-lowpop+1 # Not used currently; would be incorporated if upper and lower
harvest/escapement dynamics are incorporated
      lcv<-log(mrstd^2+1) # log CV
      modintabf[i,8,v]<-mrstd*sum(popmat.f[i+8,1,,v]) #changed from saving in column 12; removed
firstup:40 from third index
      modintabf[i,7,v]<- (mrmn*sum(popmat.f[i+8,1,,v])*exp(rnorm(1,-1*(lcv)^2)/2,sqrt(lcv))))+((1-
mrmn)*sum(popmat.f[i+8,1,,v])) # All years actual run size; changed from saving in column 11
    }

```

```

# Columns 13-18 are weir indices of escapement
for(j in 1:nweir){
  pop<-weirs[j] # which subpopulation to index?
  modintabf[i,12+j,v]<-obsmatf[i+8,2,pop,v] # Weir escapement
}

# Columns 19-32 are aerial indices of escapement
for(j in 1:nair){
  pop<-air[j] # which sub-population to index?
  modintabf[i,18+j,v]<-obsmatf[i+8,2,pop,v] # Aerial escapement
}

# Columns 33-42 are proportional return times for each year
modintabf[i,33:42,v]<-ret.timef[i,] # Proportion of run returning in week x

# columns 41,44,47,50,53,56,59 are Fishing harvest in week x
for(j in 1:nwk)
{
  modintabf[i,(40+3*j),v]<-sum(harvdyf[i+8,1:4,,j,v]) # Harvest by week across all ages and
populations and fisheries
}

# Columns 42,45,48,51,54,57,60 are effort in week x
for(j in 1:nwk)
{
  modintabf[i,(41+3*j),v]<-effort.timef[i,v] # effort by week
}

# columns 43,46,49,52,55,58,61 are Fishing regulation in week x

```

```

for(j in 1:nwk)
{
  modintabf[i,(42+3*j),v]<-1 #Fishing regulation by week (only one regulation simulated)
}

}

substock.runsf[,v]<-popmat.f[,1,,v] # Substock run dynamics
substock.escf[,v]<-popmat.f[,2,,v] # Sub-stock escapement dynamics
}

}

if(!forecast)return(list("Summary" = modintab,"SubStock"= substock.runs,"Weirs" = weirs,"Aerial" = air,
  "SiteCorr"=spatialcorr,"TrueEsc"=substock.esc))
if(forecast) return(list("Summary" = modintab,"SubStock"= substock.runs,"Weirs" = weirs,"Aerial" = air,
  "SiteCorr"=spatialcorr,"TrueEsc"=substock.esc,"SRParms"=srparmsmat,"SRSims"=modintabf,
  "SRSimRuns"=substock.runsf,"SRSimsEsc"=substock.escf))
}

#=====
# Function to generate recruits from a spawning population
#
# Parameters:
# - spawn: number of spawners
# - sig: process error in recruitment
# - a: alpha parameter
# - b: beta parameter
# - rho: temporal autocorrelation in recruitment
# - last.eps: last time step recruitment anomaly

```

```

# - anom: recruitment anomaly
# - srtype: Shape of stock-recruit curve ("ricker" or "bevholt")
#=====
rick<-function(spawn,sig,a,b,rho,last.eps,anom,srtype="ricker")
{
  eps<-rho*last.eps+sqrt(1-rho^2)*anom # generate temporal autocorrelation component

  if(srtype=="ricker"){
    lrs<-log(a)+b*spawn+eps*sig-(sig^2)/2 # Generate log(recruits per spawners)
    rec<-exp(lrs)*spawn # Generate recruits from lrs
  }

  if(srtype=="bevholt"){
    rec<-((a*spawn)/(b+spawn))*exp(eps*sig-(sig^2)/2) # Generate recruits
  }

  return(list=c(rec,eps)) # return recruits and temporal autocorrelation component
}

#=====
# Function to generate age composition from Dirichlet distribution
#
# Parameters:
# - recruits: number of recruits from brood year
# - npop: number of populations
#=====
age.comp<-function(recruits,npop)

```

```

{
  age<-rdirichlet(npop,c(.19,.38,.390,.04)*137) # generate age-composition from dirichlet
  recclasses<-round(age*recruits) # generate recruitment by age class
  return(recclasses) # return recruits by age classes
}

#=====
# Function to generate return timing
#
# Parameters:
# - npop: number of populations
#=====
returntime<-function(npop)
{
  timing<-rdirichlet(npop,c(.006,.07,.186,.286,.225,.128,.052,.025,.011,.011)*75) #generate return timing
  proportions
  return(timing) # return salmon return timing proportions
}

#=====
# Function to generate weekly fishing effort multipliers. These
# values are multiplied by the annual effort level to determine
# weekly effort. Note: not currently used in simulations
#
# Parameters:
# - npop: number of populations (set to one, as harvest affects
# all populations proportionally to their abundance)
#=====
effort<-function(npop=1)

```

```

{

eff<-rlnorm(7,-.2,.4)# Draw log-normal multiplier values for each week of harvest
eff<-eff[order(eff)][c(4,5,7,6,3,2,1)] # Order the values to emulate true harvest dynamics
return(eff)
}

#=====
# Function to generate time-series of recruitment anomalies for
# subpopulations with different correlation structures.
# Note: Currently do not use the "perfect" structure, as these
# dynamics can be produced with the "expdecay" structure
# when decaymin = 1
#
# Parameters:
# - npop: number of populations
# - synchrony: type of synchrony among populations (Use "expdecay")
# - nyr: number of years to simulate
# - procerr: sd of recruitment error
# - decaymin: minimum correlation among sub-populations
#=====
generate.anoms<-function(npop,synchrony=c("none","perfect","expdecay"),nyr,procerr,decaymin=0)
{

if(synchrony=="expdecay")
{

lindec<-(seq(1,decaymin,length=npop)) # sequence declining linearly from 1 to decaymin
ma<-matrix(0,nrow=npop,ncol=npop) # storage matrix for correlation matrix

```

```

ma[,1]<-lindec # Linear decline in column 1
for(i in 2:npop){
  ma[-c(1:(i-1)),i]<-lindec[1:(npop-(i-1))] # Fill in lower triangle
}
ma<-ma+t(ma)-diag(nrow(ma)) # fill in upper triangle -> correlation matrix
sds<-rep(procerr,npop) # standard deviations
b<-sds%*%t(sds)
spatialcorr<-b*ma # generate covariance matrix

  anomns<-rmvnorm(nyr,mean=(-1*diag(spatialcorr)/2),sigma=spatialcorr) # draw mvnorm samples from
covariance matrix

}

if(synchrony=="perfect")
{
  anomns<-rnorm(nyr,mean=(-1*procerr^2/2),sd=procerr) # generate one anomaly per year
  anomns<-matrix(anomns,nrow=nyr,ncol=npop,byrow=F) # apply this anomaly to all populations
  spatialcorr<-NA
}

if(synchrony=="none")
{
  anomns<-matrix(rnorm(nyr*npop,mean=(-1*procerr^2/2),sd=procerr),nrow=nyr,ncol=npop,byrow=F)
# draw one anomaly per population per year
  spatialcorr<-NA
}

return(list(anomns,spatialcorr)) # return anomalies matrix, and spatial covariance matrix

```

```
}
```

```
#=====
===
```

```
# Function to generate observations of escapement at weir sites
```

```
#
```

```
# Parameters:
```

```
# - escape: True escapement from operating model
```

```
# - obserr.wier: log-normal standard deviation of observation error on weir indices
```

```
# - weir: vector indicating which sites have weirs
```

```
# - npop: number of populations
```

```
#=====
===
```

```
observe.weir<-function(escape,obserr.weir,weir,npop=npop){
```

```
  err<-rnorm(npop,-obserr.weir^2/2,obserr.weir) # generate observation error in log space
```

```
  obsesc<-escape*exp(err)*weir # multiply true escapement by observation error for weir sites
```

```
  return(obsesc)
```

```
}
```

```
#=====
===
```

```
# Function to generate observations of escapement at aerial survey sites
```

```
# -expansion factor taken from Eggers et al. 2012 ADFG Escapement Report
```

```
# - aerial bias parameter values from Figure 8 of Jones et al 1998 NAJFM
```

```
# Parameters:
```

```
# - escape: True escapement from operating model
```

```
# - obserr.aerial: log-normal standard deviation of observation error on aerial indices
```

```
# - aerial: vector indicating which sites have aerial surveys
```

```
# - npop: number of populations
```

```

# - aerialbias: (T/F) are aerial indices increasingly biased at higher escapements
#=====
===
observe.aerial<-function(escape,obserr.aerial,npop,aerial,aerialbias=F){
  obsesc<-(escape/2.47)*exp(rnorm(npop,(-obserr.aerial^2/2),obserr.aerial))*aerial # generate aerial
index by adding log-normal observation error
                                # to a constant proportion of the trueescapement
  if(aerialbias) obsesc<- exp(0.59)*(escape^0.83)*exp(rnorm(npop,(-
(obserr.aerial^2)/2),obserr.aerial))*aerial # generate aerial index by adding log-normal observation error
                                # to an increasingly negatively biased
proportion of the true escapement
  return(obsesc)
}

```

Appendix 5. ADMB code for current run-reconstruction model ('R model').

```
//=====
// Converting Alaska Department of Fish and Game
// Kuskokwim River Chinook salmon R-based run-reconstruction
// model into ADMB ("original model"
//
// Model by Hamachan Hamazaki
//
// Transferred to ADMB by Tim Walsworth (tewals@uw.edu)
//=====

//=====
//DATA SECTION
//=====
DATA_SECTION
  init_int nyear; // number of years with data
  init_int nweek; // number of weeks for harvest data
  init_int nweir; // number of weir sites
  init_int nair; // number of aerial survey sites

  init_matrix testf(1,nyear,1,nweek); //Estimates of run proportion by
week

  init_matrix ceff(1,nyear,1,nweek); // Weekly effort commercial fishery
  init_matrix ccat(1,nyear,1,nweek); // Weekly catch commercial fishery
  init_matrix creg(1,nyear,1,nweek); // Weekly indicator of fishery
regulation

  init_vector inriv(1,nyear); // Annual in-river run estimate
  init_vector inriv_sd(1,nyear); // SD of annual in-river run
estimate

  init_vector tcatch(1,nyear); // Total harvest across all
fishery sectors
  init_matrix esc_w(1,nyear,1,nweir); // Weir escapement indices
  init_matrix esc_a(1,nyear,1,nair); // Aerial escapement indices

  // init_vector minesc(1,nyear); // Minimum annual escapement
  init_vector minrun(1,nyear); // Minimum annual run size
  init_vector ubrun(1,nyear); // Upper bounds for annual run
size estimation
// cout <<"Data Section Complete"<<endl;

//=====
// Parameter Section
//=====
PARAMETER_SECTION
  init_bounded_number_vector log_trun(1,nyear,minrun,ubrun,1); // log
drainage-wide run
  init_bounded_vector log_wesc(1,nweir,2,7,1); // log slope for weir
counts
```

```

    init_bounded_vector log_aesc(1,nair,3,8,1);    // log slope for aerial
counts
    init_bounded_vector log_q(1,3,-14,-5,1);    // log Catchability for
different fishery sectors
    init_bounded_vector log_rw(1,nweir,-3,5,1);  // log overdispersion for
weir counts //upper bound was 10
    init_bounded_vector log_ra(1,nair,-3,5,1);   // log overdispersion for
aerial counts //upper bound was 10

    vector t_run(1,nyear);                        // storage for untransformed total
runs
    vector wesc(1,nweir);                        // storage for untransformed weir
escapement slopes
    vector aesc(1,nair);                        // storage for untransformed aerial
escapement slopes
    vector q(1,3);                              // storage for untransformed
catchabilities
    vector rw(1,nweir);                        // storage for untransformed weir
overdispersion parameters
    vector ra(1,nair);                        // storage for untransformed aerial
overdispersion parameters
    matrix wk_est(1,nyear,1,nweek);            // storage matrix for the estimated
number of fish available for harvest each week
    number tfw;                                // likelihood for weir counts
    number tfa;                                // likelihood for aerial counts
    number tfc1;                               // likelihood for catch data gear
1
    number tfcpart1;
    number tfc2;                               // likelihood for catch data gear
2
    number tfcpart2;
    number tfc3;                               // likelihood for catch data gear
3
    number tfcpart3;
    number tft;                                // likelihood for in-river run
estimates
    vector esc(1,nyear);                       // vector of total escapement
estimates

    objective_function_value objf;
// cout <<"Parameter Section Complete"<<endl;

INITIALIZATION_SECTION
    log_trun 12.5;
    log_wesc 5.0;
    log_aesc 4.0;
    log_q -11.0;
    log_rw 1.0;
    log_ra 1.0;

// cout <<"Initialization Section Complete"<<endl;

//=====

```

```

// Procedure Section
//=====
PROCEDURE_SECTION

    objf = 0.0;

    convert_parameters_into_rates();
//cout<<"Convert Parameters Complete"<<endl;

    evaluate_obj_func();
// cout<<"Evaluate Objective Function Complete"<<endl;

RUNTIME_SECTION
    maximum_function_evaluations 200000000
    convergence_criteria 1.e-30

//=====
// Function convert_parameters_into_rates
//=====
FUNCTION convert_parameters_into_rates

    t_run=exp(log_trun);
    wesc=exp(log_wesc);
    aesc=exp(log_aesc);
    q=exp(log_q);
    rw=exp(log_rw);
    ra=exp(log_ra);

//=====
// Function evaluate_obj_func
//=====
FUNCTION evaluate_obj_func
    int i,j,k,l,ctr1,ctr2,ctr3;

    // Initialize values
    tfw= 0.0;
    tfa= 0.0;
    tft= 0.0;
    tfcpart1=0.0;
    tfc1=0;
    tfcpart2=0.0;
    tfc2=0;
    tfcpart3=0.0;
    tfc3=0;
    ctr1=0;
    ctr2=0;
    ctr3=0;

    for (i=1;i<=nyear;i++)
    {
        esc(i)=t_run(i)-tcatch(i); // calculate escapement
    }

```

```

    if(inriv(i)>0)
    {
        tft+= 0.5*sqrt((inriv(i)-(t_run(i)))/sqrt(inriv_sd(i))); //
In-River run estimate likelihood
    }

    // Weir likelihoods
    for(j=1;j<=nweir;j++)
    {
        if(esc_w(i,j)>0)
        {
            tfw+= -1*(gammln(esc_w(i,j)+rw(j))-gammln(esc_w(i,j)+1)-
gammln(rw(j))+rw(j)*log(rw(j)/(esc(i)/wesc(j)+rw(j)))+esc_w(i,j)*log((esc(
i)/wesc(j))/(esc(i)/wesc(j)+rw(j))));
        }
    }

    // Aerial likelihoods
    for(k=1;k<=nair;k++)
    {
        if(esc_a(i,k)>0)
        {
            tfa+= -1*(gammln(esc_a(i,k)+ra(k))-gammln(esc_a(i,k)+1)-
gammln(ra(k))+ra(k)*log(ra(k)/(esc(i)/aesc(k)+ra(k)))+esc_a(i,k)*log((esc(
i)/aesc(k))/(esc(i)/aesc(k)+ra(k))));
        }
    }

    // Harvest and effort likelihoods
    for(l=1;l<=nweek;l++)
    {

        wk_est(i,l) = t_run(i)*testf(i,l);
        if(ccat(i,l)>0)
        {

            if((1-ccat(i,l)/wk_est(i,l))>0)
            {

                if(creg(i,l)==1){
                    tfcpart1+= square(log(ceff(i,l))-log((-log(1-
ccat(i,l)/wk_est(i,l)))/q(1))));
                    ctr1++;
                }

                if(creg(i,l)==2){
                    tfcpart2+= square(log(ceff(i,l))-log((-log(1-
ccat(i,l)/wk_est(i,l)))/q(2))));
                    ctr2++;
                }
            }
        }
    }

```

```

        if(creg(i,1)==3){
            tfcpart3+= square(log(ceff(i,1))-log((-log(1-
ccat(i,1)/wk_est(i,1)))/q(3))));
            ctr3++;

        }

    }

}

}

}

tfc1= 0.5*ctr1*log(tfcpart1);
tfc2= 0.5*ctr2*log(tfcpart2);
tfc3= 0.5*ctr3*log(tfcpart3);

objf+= tft+tfw+tfa+tfc1+tfc2+tfc3; // Objective Function

//=====
=====
// Report Section
//=====
=====
REPORT_SECTION

    report<<"#Total Run"<<endl;
    report<< t_run << endl;
    report<<"#ObjFunc"<<endl;
    report<< objf << endl;
    report<<"#rw"<<endl;
    report<< rw << endl;
    report<<"#ra"<<endl;
    report<< ra << endl;
    report<< "#q" << endl;
    report<< q << endl;
    report<< "#wesc" << endl;
    report<< wesc << endl;
    report<< "#aesc" << endl;
    report<< aesc << endl;
    report<<"#tcatch"<<endl;
    report<< tcatch<<endl;
    report<<"#Total Escapement"<<endl;
    report<< esc << endl;

//=====
=====
// Globals Section
//=====
=====

```

```

GLOBALS_SECTION
#include <df1b2fun.h>
#include <math.h>
#include <time.h>
#include <statsLib.h>
#include <adrndeff.h>
#include <admodel.h>
time_t start,finish;
long hour,minute,second;
double elapsed_time;

TOP_OF_MAIN_SECTION
arrmblsize = 100000000;
gradient_structure::set_MAX_NVAR_OFFSET(30000000);
gradient_structure::set_GRADSTACK_BUFFER_SIZE(3000000);
gradient_structure::set_CMPDIF_BUFFER_SIZE(100000000);
time(&start);

FINAL_SECTION
// Output summary stuff
time(&finish);
elapsed_time = difftime(finish,start);
hour = long(elapsed_time)/3600;
minute = long(elapsed_time)%3600/60;
second = (long(elapsed_time)%3600)%60;
cout << endl << endl << "Starting time: " << ctime(&start);
cout << "Finishing time: " << ctime(&finish);
cout << "This run took: " << hour << " hours, " << minute << " minutes,
" << second << " seconds." << endl << endl;

```

Appendix 6. AYK SSI EXPERT PANEL MEMBER BIOGRAPHIES

Daniel E. Schindler (Chair)

deschind@uw.edu

Dr. Daniel Schindler is a Professor in the School of Aquatic and Fishery Sciences at the University of Washington where he has been on the faculty since 1997. He is a principal investigator of the UW Alaska Salmon Program which has studied the ecology of salmon and their watersheds in western Alaska since 1946. His research is focused on freshwater ecosystems and their watersheds, addressing questions ranging from understanding basic ecological and evolutionary processes, to the effects of climate change, watershed development and fisheries on natural resources. In Bristol Bay his research has quantified why the vast and intact watersheds have produced reliable and sustainable runs of salmon, despite high harvest rates, warming climate, and few management interventions. He received a B. Sc. with Honours from the University of British Columbia, and a M.S. and Ph.D. from the University of Wisconsin-Madison. He has published over 180 peer-reviewed scientific papers and serves as an editor of the journals *Ecology*, *Ecosystems*, and *FACETS*. He spends over 3 months of the year in the field in western Alaska, and has provided professional service to a wide variety of governmental and non-governmental organizations.

Milo Adkison

mdadkison@alaska.edu

Dr. Milo Adkison is a Professor of Fisheries and Chair of the Fisheries Department at the College of Fisheries and Ocean Sciences, University of Alaska Fairbanks, where he's worked since 1997. He's also worked for the Biological Resources Division of USGS and as crew on a Bristol Bay gillnetter. The research of Dr. Adkison and his students typically involves applying quantitative methods to aspects of Pacific salmon biology and fisheries management. Recent studies have focused on nutrients imported from the ocean by salmon, survival of Chinook salmon, bycatch of salmon in groundfish fisheries, and improving stock assessment models.

Randall M. Peterman

peterman@sfu.ca

Dr. Randall M. Peterman retired in 2012 from the School of Resource and Environmental Management at Simon Fraser University (Burnaby, British Columbia, Canada), where he held the position of Professor and Canada Research Chair in Fisheries Risk Assessment and Management. For over 40 years, his award-winning research has focused on quantitative methods to improve understanding of population dynamics of Pacific salmon and their management, particularly in the presence of uncertainties and risks. He has served on more than 25 professional committees, including (1) the Pacific Salmon Commission's 2016 Independent Expert Panel review of Chinook salmon forecasting methods, (2) the 2013 Independent Advisory Panel on the decline of southern British Columbia Chinook salmon, (3) the Pacific Salmon Commission's 2010 Expert Panel on investigating the causes of the decline in productivity of Fraser River

Sockeye, (4) the 2008 Skeena Independent Science Review Panel, and (5) the Royal Society of Canada's 2012 Expert Panel on "Sustaining Canada's Marine Biodiversity: Responding to the Challenges Posed by Climate Change, Fisheries, and Aquaculture".

André E. Punt

aepunt@uw.edu

Dr. André E. Punt is a Professor in the School of Aquatic and Fishery Sciences at the University Washington, Seattle, USA and the currently the Director of the School. He received his B.Sc., M.Sc. and Ph.D. in Applied Mathematics at the University of Cape Town, South Africa. The research undertaken by Dr. Punt and his research group relates broadly to the development and application of fisheries stock assessment techniques, bioeconomic modelling, and the evaluation of the performance of stock assessment methods and harvest control rules using the Management Strategy Evaluation approach. Dr. Punt has conducted stock assessments for a wide range of species, ranging from anchovies and sardines, to groundfish, tunas, and cetaceans. He has published over 300 papers in the peer-reviewed literature, along with over 400 technical reports. Dr. Punt is currently a member of the Scientific and Statistical Committee of the Pacific Fishery Management Council, the advisory committee for Center for the Advancement of Population Assessment Methodology, the Crab Plan Team of the North Pacific Fishery Management Council, and the Scientific Committee of the International Whaling Commission.

Timothy Walsworth (analyst)

timothy.walsworth@usu.edu

Dr. Timothy Walsworth is a post-doctoral researcher at Utah State University. He received his B.S. in Zoology from the University of Wisconsin-Madison, his M.S. in Ecology from Utah State University, and his Ph.D. in Aquatic and Fishery Sciences from the University of Washington. His research focuses on using quantitative modeling approaches to examine the potential response of biological communities to environmental and management changes, thus allowing stakeholders and managers to consider potential trade-offs among alternative approaches to management of aquatic resources. He has conducted research in diverse ecological systems and issues, from highly altered desert streams to pristine Alaskan watersheds, and from endangered species population dynamics to commercial fishery bioeconomic models.