March 2026

Spring-Run Juvenile Production Estimate Independent Peer Review

A report to the Delta Science Program

Prepared by

Charles Chamberlain (Lead Author), 100 Dog Fluvial
Greg Ruggerone, Ph.D. (Panel Chair) Salmon Science
Kelly Vasbinder, Ph.D. University of California, Santa Cruz
Julian Olden, Ph.D. Procrustes Environmental Consults
Carl Schwarz, Ph.D. StatMathComp Consulting by Schwarz Inc.

Delta Science Program
DELTA STEWARDSHIP COUNCIL

Table of Contents

Abbreviations

Term	Definition
AIC	Aikake Information Criteria
BHM	Bayesian Hierarchical Model
BT-SPAS	Bayesian Temporally Stratified Population Analysis System
BUGS	Bayesian inference using Gibbs sampling
CDFW	California Department of Fish and Wildlife
CJS	Cormack-Jolly-Seber
CV	coefficient of variation
CWT	coded wire tag
Delta	Sacramento–San Joaquin River Delta
DWR	California Department of Water Resources
JAGS	Just Another Gibbs Sampler
JPE	juvenile production estimate
LAD	length-at-date
LOOCV	leave-one-out cross validation
MCMC	Markov Chain Monte Carlo

mm	millimeter
MT	Modeling Team
PIT tags	Passive integrated transponder tags
PLAD	probabilistic length at date
PPO	prior-posterior overlap
PPP	Posterior predictive plot
RBDD	Red Bluff Diversion Dam (rotary screw trap site)
Rhat	Potential Scale Reduction Factor (Gelman-Rubin statistic)
RST	rotary screw trap
SPAS	Stratified Population Analysis System
SRJPE	spring-run juvenile production estimate
USFWS	U.S. Fish and Wildlife Service

Executive Summary

As part of California Endangered Species Act and federal Endangered Species Act consultations on the long-term operations of the state and federal water projects, the California Department of Water Resources (DWR) and the U.S. Bureau of Reclamation agreed to support development and implementation of an approach for forecasting an annual spring-run Chinook salmon (*Oncorhynchus tshawytscha*) juvenile production estimate (SRJPE) for the Sacramento River and its tributaries upstream of the Sacramento–San Joaquin Delta (Delta). The Juvenile Production Estimate Science Program was initiated through these consultations with multiple teams made up of various disciplines and responsibilities. The Modeling Team (MT) of this program was tasked to generate the SRJPE model and draft a report. The production of an annual SRJPE forecast is intended to support the development of new measures to minimize the loss of spring-run Chinook salmon caused by water operations and is also intended to contribute to the development of a spring-run life cycle model to support improved management of spring-run Chinook salmon.

In December 2025, the MT produced a draft report, *Spring-run Chinook Salmon Juvenile Production Estimate Modeling Approach* (Modeling Team Draft Report 2025), containing nine chapters that described an integrated model and multiple submodels designed to:

1. Develop, evaluate, and recommend a suite of potential approaches to estimate the abundance and timing of SRJPE in the Sacramento River and its tributaries upstream of the Delta, and
2. Support the potential development of a life cycle model and minimization measure without knowing the form the life cycle model and/or minimization measure would take.

The MT continues to develop the SRJPE, and the current report is not the endpoint of their effort. The MT described the forecasted SRJPE estimate as a demonstration of how the integrated model works and stated that it is not intended to provide a definitive in-season forecast at this time.

This Review Panel was coordinated by the Delta Science Program to review the current state of the modeling effort to ensure that the project is on track to meet its objectives. The MT briefly mentioned some new information and preliminary findings at the February 24, 2026 *Spring-Run Juvenile Production Estimate Peer Review Public Meeting*. This review focuses on information in the Modeling Team Draft Report (2025) unless otherwise noted. It responds to the scope and charge questions presented to the Review Panel by the Delta Science Program and offers a detailed review of each chapter in the MT report (i.e., models, submodels and integration framework). The review represents the shared professional opinion and recommendations of the five-member Review Panel, which is independent of the MT.

The proposed SRJPE provides a structured approach to estimating the annual number of spring-run juveniles entering the Delta while also quantifying much of the associated uncertainty. By integrating multiple submodels, the SRJPE attempts to maximize the use of available data from tributary rotary screw traps (RST) and forecast survival to Delta entry. The goal is to offer a repeatable approach that produces more precise juvenile production estimates than those from individual submodels alone. Further, this suggested approach reduces subjective decisions, such as when to switch between pre-season and in-season forecasting.

The SRJPE approach is innovative and uses a variety of existing datasets. The SRJPE can approximate the true number and timing of spring-run juveniles entering the Delta within its current limitations, providing a useful preliminary estimate with associated uncertainty. The framework does provide credible intervals to quantify uncertainty, but these intervals may not fully capture all sources of bias or uncertainty. The Review Panel offers numerous suggestions in this Review Report to strengthen the SRJPE which include out-of-sample testing, verification of whether large acoustically-tagged smolts represent much smaller fry when estimating survival and travel time from RSTs to the Delta, validation that hatchery fish represent wild fish when measuring capture-efficiency at the RST and survival and travel times in the mainstem, expansion of the probabilistic length at date (PLAD) assignments to other tributaries (e.g., Yuba River), usage of additional tributary data (e.g., Feather River), improving covariate forecasting, and performing additional model refinements to better align juvenile production estimates with the true annual juvenile abundance while also accurately quantifying the error around these estimates.

A breakdown of total uncertainty into contributions from different sources, such as RST measurement error, survival model uncertainty, and covariate variability, is recommended to focus subsequent efforts on improving the models/submodels and data collection/monitoring. Presently, the greatest source of uncertainty in the SRJPE appears to stem from the survival and travel time of salmon as they migrate from tributary RSTs to the Delta entry. This uncertainty does not incorporate potential bias associated with the extrapolation of large acoustically-tagged salmon to mimic the survival and migration of salmon fry, which represent the majority of the SRJPE. Further analysis would help identify where the most important improvements could be made. If a time series of tributary-specific abundances of spring-run juveniles entering the Delta becomes available through this modeling effort, the MT could compare these estimates with adult returns to the watersheds to evaluate whether the survival values seem reasonable and consistent across populations.

The MT was tasked with generating the SRJPE and informing life cycle models without knowing specific management actions that might be used to minimize adverse effects on spring-run juveniles. Knowing about potential management actions that might be undertaken could reduce uncertainty by allowing specific parts of the SRJPE to be targeted. It could also reduce the level of monitoring required for the SRJPE. The Review Panel recommends that guidance on the most likely management actions be developed to help the MT focus on improving the modeling effort.

A notable mainstem trapping location not included in the SRJPE model is the U.S. Fish and Wildlife Service (USFWS) site at Red Bluff Diversion Dam (RBDD). It employs as many as five rotary screw traps to achieve higher capture probabilities than the Tisdale or Knights Landing sites. The Review Panel recommends an investigation into using this valuable additional data.

The MT Report should evaluate various flood control structures along the Sacramento River from the Sutter Basin to Delta entry or the Yolo Bypass and the flows that entrain, or draw in, juvenile salmon. At multiple points along the river, structures spill floodwaters (and juvenile salmon) to large flood-bypass or floodplain areas that provide growth and potential survival benefits to the entrained juvenile salmon. Those flood-control features are overtopped at various discrete flow levels. Assessing juvenile migratory speed, growth, or survival at those discrete flow values may yield more information than an assessment based on categorical water year type or peak flows alone. Also, the MT should consider that some of these flood control structures route fish around the mainstem RST sites currently utilized in the SRJPE model, but not the Red Bluff site. Additional monitoring may be required to assess the effects of flood flow on the timing, growth, and survival of spring-run juveniles entering the Delta.

The Review Panel also provides suggestions on potential modifications to the current monitoring program to improve forecasts and resolve some of the issues raised in our review. The Review Panel points to other potential datasets along with emerging technologies that may develop in useful tools for SRJPE including acoustic tags that are being developed to tag much smaller fish, passive integrated transponder (PIT) tagging applications, hydroacoustics, eDNA, investigating the use of Chipps Island Trawl data, use of available coded wire tag (CWT) data (the MT mentioned CWT analysis currently underway) and application of Parental-Based Tagging (especially in the Feather River with its hatchery production).

Acoustic tag technology continues to improve, leading to much smaller tags than those currently used in the Sacramento River. The Pacific Northwest National Laboratory has developed a small fish shad tag weighing only 0.05 grams that has been used on fish as small as 50 millimeters. The Review Panel believes this new tag may offer the best opportunity to improve survival and migration rates in the mainstem river.

The MT and Review Panel found that the relatively large Feather and Yuba rivers currently have less realistic spring-run Chinook salmon values than the other populations, in part due to issues with run-type identification, hatchery production, hybridization, and limited availability of concomitant adult and juvenile monitoring data. Weekly estimates of juvenile spring-run abundances at RST sites on the Feather River were not available in the Modeling Team Draft Report (2025) due to complex run type identification issues. Work has been ongoing to apply the PLAD model to data from the Feather River to identify run type. Identification of Chinook run types on the Feather River is more challenging than at other RST sites due to a more complicated run structure, the release of large numbers of hatchery-origin spring-run and fall-run juveniles upstream of most RST sites (approximately 2 million spring-run and approximately 2.3 million fall-run juveniles released per year), spring/fall-run hybrids produced by adults interbreeding on the spawning grounds, and the unintentional inclusion of untagged spring-run Chinook salmon in the fall-run hatchery broodstock. These issues affect the accuracy of PLAD models, which assign run types. The MT indicated at the February 24, 2026 *Spring-Run Juvenile Production Estimate Peer Review Public Meeting* that they are making progress toward addressing these Feather River SRJPE issues. The Review Panel has not examined new analyses since publication of the 2025 report.

Given the Feather River's large production of naturally-spawned spring-run Chinook salmon, complicated as their run structure may be, these issues could have a substantial influence on the accuracy of the Sacramento River-wide SRJPE. Consequently, the Review Panel concurs with the MT and recommends:

1. Modeling heterozygotes explicitly,
2. using Eye Riffle RST data for its reduced hatchery influence and co-location of genetic testing,
3. and separation of the Feather River spring-run juvenile production estimate from the valley-wide SRJPE.

Three populations are considered independent, Butte, Deer, and Mill creeks, meaning they are self-sustaining and not dependent on strays from other watersheds, unlike the dependent populations (e.g., Battle and Clear creeks). Including abundance values from the Feather and Yuba rivers may skew the total SRJPE and the potential management actions. Thus, abundance values from individual populations should continue to be evaluated, especially the independent populations that may have the greatest value for conservation.

The MT documents could benefit from a discussion of the potential for future data loss at small or large scales and the impact of those losses on data quality or the integrity of the SRJPE output. There are examples of short-term data loss on the order of days or weeks in the existing historical datasets including harsh weather or flows high enough to make juvenile trapping programs pause for safety, limited spawning ground surveys for storms or fires, reduced trapping effort when programs are occasionally understaffed (shorter trapping seasons or trapping less than 7 days a week), or government shutdowns that require affected agencies to cease work temporarily. Longer term loss of entire program components is a real risk too, given funding uncertainties. Contingency planning or anticipatory adjustments are not fully addressed in the documents.

Additional statistical analyses may provide contingencies for unplanned losses in monitoring operations. For example, the MT efforts to reconstruct or simulate missing portions of the data portfolio, such as applying PLAD results to years prior to genetic testing, are adjustments that could prove useful for addressing future missing components. Correlations in juvenile abundances among tributaries could be explored for the utility of using neighboring tributary data to estimate juvenile production if an entire tributary monitoring program is shut down for loss of staff or funding. High correlations can be used to impute values when data from a tributary are missing in a year (e.g., forest fire causes suspension of a monitoring site). Or they may imply a shared environmental factor across all tributaries that could serve as a common covariate. Finally, a more complex Bayesian model could be fit to use these cross-correlations to reduce the uncertainty in tributary-specific estimates, similar to the effect of “averaging” across tributaries. The mainstem RST site at Red Bluff Diversion Dam should also be explored for its utility in estimating production and timing from upstream tributaries as potential insurance against the loss of one or more of those monitoring programs.

Overall, the Review Panel found the model’s technical aspects generally appropriate. Nevertheless, the Review Panel provides several technical observations and suggestions for model improvement. These are detailed in the Review of Material and Supplemental Documents.

While this Review Panel Report offers numerous comments and recommendations for the current state of the SRJPE, the Panel commends the MT for a massive amount of work and innovation toward applying data from multiple monitoring programs — many of which were initiated and structured long ago for purposes other than a Sacramento River-wide SRJPE. The MT assembled a viable SRJPE approach that both leans on historical data and informs future monitoring to strengthen the SRJPE. The MT was hindered by externally imposed limitations, primarily that it was directed to assemble the SRJPE without presuming which conservation or take-minimization measures the model might inform. It's the opinion of the Review Panel that the MT could be more effective if informed of the decision points they are modeling toward.

1. Introduction

As part of California Endangered Species Act and federal Endangered Species Act consultations on the long-term operations of the state and federal water projects, the California Department of Water Resources (DWR) and the U.S. Bureau of Reclamation agreed to support development and implementation of an approach for forecasting an annual spring-run Chinook salmon (*Oncorhynchus tshawytscha*) juvenile production estimate (SRJPE) for the Sacramento River and its tributaries upstream of the Sacramento–San Joaquin Delta (Delta) (Modeling Team Draft Report 2025). The production of an annual SRJPE forecast is intended to support the development of new potential measures to minimize the loss of spring-run caused by water operations, and to contribute to the development of a spring-run life cycle model to support improved management.

In response to these issues, the JPE Core Team and the MT were tasked with an effort to address these overarching goals:

1. Develop, evaluate, and recommend a suite of potential approaches to estimate the abundance and timing of spring-run Chinook salmon juveniles in the Sacramento River and its tributaries upstream of the Delta, and
2. Support the potential development of a life cycle model and minimization measure.

Importantly, the JPE Core and Modeling Teams did not know what form the life cycle model and/or minimization measure would take. Additional refinement and improvement of the SRJPE could be made when this information becomes available.

In December 2025, the MT produced a draft report containing nine chapters that described an integrated model and multiple submodels designed to achieve these goals[1]. Chapters 1 and 2 provide background information, Chapters 4 to 9 describe submodels, and Chapter 3 integrates information

---

[1] Modeling Team documents available at deltacouncil.ca.gov/delta-science-program/spring-run-juvenile-production-estimate-independent-peer-review See “Review Documents”

from the submodels to estimate the abundance and timing of spring-run juvenile spring Chinook salmon entering the Delta (i.e., SRJPE).

The modeling effort is ongoing. The MT briefly mentioned some new information at the February 24, 2026 *Spring-Run Juvenile Production Estimate Peer Review Public Meeting*, where the Review Panel presented their preliminary findings. This review focuses on information provided in the December draft reports unless otherwise noted.

Charge to the Independent Review Panel

Scope of Review

1. Provide a technical evaluation of the scientific approach developed for calculating an annual SRJPE including its scientific rigor and make recommendations to improve the modeling approach.
2. Evaluate the ability of the SRJPE to describe characteristics (e.g., abundance, source, timing) of the juvenile spring-run population entering the Delta.
3. Evaluate the ability of the SRJPE to inform a spring-run Life Cycle Model
4. Provide clear and implementable recommendations for high level changes as well as specific recommendations to improve the ability of the JPE to achieve the intended uses described above.

Charge Questions

1. To what extent does the proposed SRJPE provide an annual estimate of the number of spring-run juveniles entering the Delta that reasonably approximates the true number including the error associated with that estimate?
    a. Are there additional ways to evaluate the ability of the SRJPE to approximate the true number of spring-run juveniles entering the Delta?
2. How could the proposed SRJPE plan make more effective use of historical and current data?
    a. Are there additional data/monitoring needs or modifications to existing monitoring approaches that could be considered?

3. Does the proposed SRJPE plan include adequate contingencies for unplanned events such as the loss of adult and juvenile monitoring program data due to extreme events?

4. Are there additional ways to reduce uncertainties (e.g., changes to modeling or monitoring) in the SRJPE model as it relates to:
    a. Accounting for the diverse life history strategies exhibited by the Central Valley spring-run population (e.g., How can the SRJPE model incorporate yearlings?).
    b. The proposed periodic model updates and whether they provide sufficient opportunities for improvement as new monitoring or science (e.g., Parentage-Based Tagging) become available?
    c. Any other observed concerns or deficiencies in the modeling approach?

5. Are there any suggestions for how to deal with the challenges described for modeling Feather River natural production (described in the In-season Outmigrant Report), which are caused by more complicated run structure than other tributaries and by large releases from the Feather River Hatchery, including unmarked fry release groups?

Report Outline

This report begins with the Review Panel's responses to the Charge Questions. These responses are designed to be synoptic, whereas more details, including additional recommendations, are presented in the review of each of the nine chapters that follows. Minor comments specific to each chapter are provided in Appendix A .

Panel Members

The Delta Science Program convened a Panel of experts to perform the review.[2] The Panel consisted of: Dr. Greg Ruggerone (Panel Chair), Charles Chamberlain (Lead Author), Dr. Kelly Vasbinder, Dr. Julian Olden, and Dr. Carl Schwarz. Members were selected to ensure that the Panel had sufficient expertise in the technical areas needed to address the Scope and Charge questions (see above).

---

[2] Biographies of the Review Panel members are posted at deltacouncil.ca.gov/delta-science-program/spring-run-juvenile-production-estimate-independent-peer-review

While this Review Panel Report offers numerous comments and recommendations for the current state of the SRJPE, the Panel commends the MT for their work and innovation toward applying data from multiple monitoring programs — many of which were initiated and structured long ago for purposes other than a Sacramento River-wide SRJPE. The MT assembled a viable SRJPE approach that both leans on historical data and informs future monitoring to strengthen the SRJPE. The MT was hindered by externally imposed limitations, primarily that it was directed to assemble the SRJPE without presuming which conservation or take-minimization measures the model might inform. It's the opinion of the Review Panel that the MT's task can be simplified and the efficacy of the SRJPE can improve dramatically when the MT is informed of the decision points they are modeling for.

Review Panel Response to Charge Questions

Question 1

To what extent does the proposed SRJPE provide an annual estimate of the number of spring-run juveniles entering the Delta that reasonably approximates the true number including the error associated with that estimate?

The proposed SRJPE provides a structured approach to estimating the annual number of spring-run juveniles entering the Delta and quantifying much of the associated uncertainty. By integrating multiple submodels, the SRJPE attempts to maximize the use of available data from tributary RSTs and forecast survival to Delta entry. The goal is to offer a repeatable approach that produces juvenile production estimates more precise than those using individual submodels alone while attempting to reduce subjective decisions, such as when to switch between pre-season and in-season forecasting.

The current SRJPE estimates should be considered preliminary, as they do not yet include all tributaries (e.g., Feather River), and key components, such as PLAD-based run type assignments at some tributaries (e.g., Yuba River), are still being refined. Furthermore, in Chapter 3, the MT described the forecasted SRJPE as a demonstration of how the integrated model works. It stated that it is not intended to provide a definitive in-season forecast at this time. Predictions using out-of-sample data from SRJPE and its submodels still need to be conducted. Substantial uncertainty remains, particularly regarding survival and travel times from tributary RSTs to the Delta, potential biases in mainstem RST catch data, and covariate forecasts for environmental conditions. Each of these sources of uncertainty are discussed below. Additional details are provided in chapter-specific comments.

Survival and travel times from tributary RSTs to the Delta

The model currently relies upon the extrapolation of survival and travel time findings from large acoustically-tagged spring- and fall-run smolts (median size 89 millimeters due to tag burden issues; often hatchery fish) to much smaller spring-run juveniles (approximately 40 millimeters) that represent the majority of the overall population. The large acoustically-tagged fish are released in spring after most of the spring-run migration has passed downstream. Thus, the size and timing of the tagged fish that provide survival and migration rate information differ from most emigrating spring-run juveniles. Extrapolations of survival and travel time to the smaller fry are generated. The estimated uncertainty in survival from the RST to Delta entry is very high, even though the median values for 5 of 6 tributaries are somewhat similar (this report Figure 1). However, these values do not account for the bias and uncertainty associated with the assumption that the tagged juveniles adequately represent most spring-run juveniles during migration. The MT is well-aware of this issue. This issue is further discussed below.

Potential biases in mainstem RST catch data

The reported SRJPE predictions in Chapter 3 highlight notable discrepancies between tributary-based forecasts and observed passage timing at mainstem RST sites like Tisdale and Knights Landing. These discrepancies are likely due in part to low and variable trap efficiencies, which reduce confidence in mainstem catch-based estimates. The time-varying nature of trap efficiencies is especially problematic. For example, environmental conditions such as declining turbidity later in the migration season may lead to increased trap-avoidance behavior and result in decreased trap efficiency that could cause underestimations of later-season passage. Additionally, run misclassifications in the PLAD model can bias the proportion of spring-run juveniles assigned to mainstem sites, particularly for populations from Mill and Deer creeks, which grow more slowly and migrate later. Together, these factors can truncate observed passage timing, making mainstem data an imperfect reference for validating SRJPE predictions at this time.

Covariate forecasts for environmental conditions

SRJPE relies on the inclusion of model covariates such as water temperature, flow, and other environmental factors to reduce model bias and uncertainty of abundance and survival estimates. Currently, forecasts use historical covariate data or annual averages, which may not capture the week-to-week variation that influences juvenile outmigration timing and survival. In many cases, it may be possible to discretize environmental variables (e.g., low, moderate, high flows or temperature bands) to generate scenario-based predictions; however, more continuous hydrographs simulated from discharge data would be optimal. Without accurate covariate forecasts, model predictions may fail to reflect extreme or unusual conditions in the forecast year, increasing uncertainty in both juvenile abundance and timing estimates. Improving covariate forecasting and integrating these forecasts into SRJPE are critical to producing more realistic and reliable Delta entry estimates. It is recommended to test alternative survival model structures and covariate selections to quantify how assumptions about survival drivers influence predictions. For example, compare survival estimates under models with different covariates (flow, temperature, size-at-date) to evaluate how these choices propagate uncertainty into Delta entry abundance forecasts.

Summary

SRJPE can approximate the true number of spring-run juveniles entering the Delta within its current limitations, providing a useful preliminary estimate with associated uncertainty. The framework does provide credible intervals to quantify uncertainty, but these intervals do not fully capture all sources of bias, such as those associated with using large acoustically tagged smolts or with missing trap-efficiency data. Out-of-sample testing, verification of whether large tagged smolts represent much smaller fry, model refinement, inclusion of additional tributary data, and improved covariate forecasting are needed to better align juvenile production estimates with the true annual juvenile abundance and to quantify the error around these estimates more accurately.

Additionally, deconstruction of total uncertainty into contributions from different sources, such as RST measurement error, survival model uncertainty, and covariate variability, is recommended. This will help managers understand which components contribute most to error and guide efforts to reduce uncertainty through additional data collection or model improvements. The MT is aware of these issues and considers the SRJPE a work in progress.

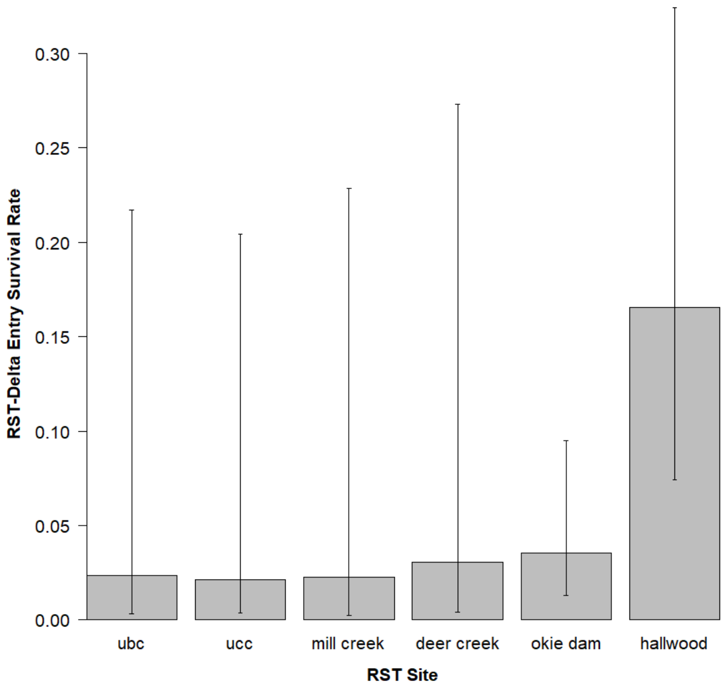


Figure 1. From the Modeling Team Draft Report (2025) Chapter 3 Figure 9. Median and 80% Credible Intervals of Survival Rate. Median and 80% credible intervals of the survival from RST sites to Delta entry used in SRJPE. These forecasts assume that peak flows during outmigration were equal to the averages across release groups from the upper Sacramento (Battle, upper Clear, Mill, Deer creeks), Butte Creek, and the Yuba River.

Question 1a.

Are there additional ways to evaluate the ability of the SRJPE to approximate the true number of spring-run juveniles entering the Delta?

The SRJPE approach is, in many ways, comprehensive in motivation but challenging to parameterize accurately in practice. The MT suggested several recommendations to improve the accuracy and reduce the uncertainty of the SRJPE. Implementing these recommendations, especially those listed in Chapter 3 and in our Chapter comments below, is the best approach for reducing bias and uncertainty in SRJPE. Below, we identify some new approaches to potentially improve the SRJPE. Each approach has complications and therefore requires additional thought before implementation.

Verify use of large acoustically-tagged smolts

Verifying the applicability of the large acoustically-tagged smolts to the earlier-timed and much smaller Chinook salmon fry is important for improving upon the accuracy of juvenile survival and migration rate in the mainstem river. Most spring-run juveniles emigrate to the Delta as fry (approximately 40 millimeters), which is much smaller than the acoustically-tagged smolts (median 89 millimeters). Most fry emigrated weeks before the release of most tagged spring- and fall-run smolts, many of which were large hatchery salmon that might have different behavior compared with natural salmon (e.g., Chen et al. 2025). Larger, later migrating smolts may be actively migrating downstream, whereas movements of fry are more likely to be passive. Chinook salmon fry are more likely to inhabit the river margin to feed and rear while avoiding predators, though some are carried downstream by the current, especially if rearing habitat is unavailable (Moyle et al. 2015). Additionally, some fry and smolts may be carried into floodplain habitats during high-water events where growth rates are higher (Sommer et al. 2001; Jeffres et al. 2008), migration rates are slowed, and survival is likely increased (Moyle et al. 2015). Given the expected differences in the behavior and migration date of 40-millimeter fry compared with approximately 89-millimeter smolts, we expect fry survival and migration rates to differ from those of the larger tagged fish. The MT attempted to address this issue by extrapolating to juvenile sizes that were much smaller than those of the tagged fish, but the applicability of this extrapolation is unknown, and it is not yet fully accounted for in the model's uncertainty estimates. Below, we identify a few new techniques that might be used to help verify the applicability of the large acoustically-tagged smolts currently used in the SRJPE.

Use smaller experimental acoustic tags

Acoustic tag technology continues to be improved, leading to much smaller tags than those currently used in the Sacramento River. Daniel Deng at the Pacific Northwest National Laboratory (zhiqun.deng@pnnl.gov) has developed a small fish "shad tag" that is only 0.05 grams (7.6 millimeters long by 2.0 millimeters in diameter). These tags are being considered for Delta Smelt (*Hypomesus transpacificus*), and laboratory tests have been conducted with smelt as small as 50 millimeters (Deters et al. 2024). The detection range depends on the location, ranging from about 70 meters to 120 meters in fresh water. Tag life at 5-second pulse rate interval is 49 days with Gen-2 microbattery. The existing acoustic tag arrays could probably be used to detect salmon containing these tags.

Evaluate Passive-Integrated Transponder (PIT) tags

PIT tags are widely used in the Columbia River basin, including the estuary, and might be used to estimate survival and travel time from release to detection near the Delta. Micro-PIT tags can be injected into salmon as small as 45 millimeters, and they are much less expensive than acoustic tags. But micro-PIT tags must be within about 11 centimeters of the antennae to be detected. This method would require deployment of large funnel nets in the mainstem river that would sample a relatively large volume of water while funneling juvenile salmon into an open cod end with PIT tag detectors. In the Columbia River estuary, a two-boat trawl is used to sample PIT tagged salmon. Their net is 91 meters wide, and they sampled the estuary for 655 hours in 2024. Details of the PIT tag methodology and findings are available (Vinarcsik et al. 2025).

Examine the marine survival of spring-run juveniles

If a time series of population-specific abundances of spring-run juveniles entering the Delta becomes available through this modeling effort, the MT could compare these estimates with adult returns to the watersheds to evaluate whether the survival values seem reasonable and consistent across populations. This approach is analogous to that used by the MT to examine salmon abundances at the RST (e.g., fry per spawner).

Enumerate salmon abundance using hydroacoustics

Hydroacoustics have been used to enumerate salmon smolts in the Columbia River and in other rivers for many years. However, those smolts are typically bigger than the majority of spring-run juvenile Chinook salmon in the Sacramento River. The 40 millimeters salmon fry would have a target strength of about -50 decibels and could be detected if background noise was minimal (C. Sullivan, Innovasea, personal communication, March 2026). But the variety of fish lengths and sampling required to identify species and run type would likely be problematic.

Population specific estimates

Some population estimates are more accurate and precise than others. For example, the relatively large Feather and Yuba rivers currently have less realistic values than the other populations, in part due to issues with run-type identification, hatchery production, and hybridization. The MT is attempting to improve the production estimates for these rivers, but the Review Panel suspects abundance values of these populations will not be as accurate as for the other populations. Additionally, three populations are considered independent (Butte, Deer, and Mill creeks), meaning they are not dependent on strays from other watersheds, unlike the dependent populations (e.g., Battle and Clear creeks). Including less-accurate abundance values from the Feather and Yuba rivers may skew the total SRJPE and the potential management actions. Thus, abundance values from individual populations should continue to be evaluated, especially the independent self-sustaining populations that may have the greatest value for conservation. Furthermore, if the actual survival from tributary RST to Delta entry is similar across most populations, as suggested in Figure 1, then the sum of population estimates from each tributary might suffice, depending on the management actions and the need to account for variability in survival and travel times in the mainstem river.

Emerging eDNA techniques

Environmental DNA techniques are being explored for their utility in quantitatively estimating fish populations or densities. As the capabilities of eDNA tools continue to develop, they may someday be useful for estimating the abundance of spring-run (Lacoursiere-Roussel et al. 2016; Spear et al. 2021; Tsuji et al. 2022).

Alternative datasets

There are other potential datasets to explore such as the Chipps Island trawl if adequate genetic samples are collected to identify run type. Length-at-date (LAD) criteria have been employed here in the past, but some dual sampling showed that many individuals classified as spring-run based on LAD were genetically fall-run (Brandes et al. 2021). Continued genetic efforts may make previous data usable if the results are sufficiently consistent to warrant applying probabilities based on length and/or date.

Question 2

How could the proposed SRJPE plan make more effective use of historical and current data?

The MT relied on a wide range of historical and current data to estimate the SRJPE. Below we highlight some potential ways in which the SRJPE plan may make better use of those data.

Consider the potential effect of hatchery releases on the migration rate of spring-run juveniles

Some evidence suggests that wild juvenile salmon may be stimulated to migrate downstream by large releases of hatchery salmon, leading to less favorable rearing by the wild salmon (Kururvilla et al. 2024). The MT could potentially test the pied-piper effect (Kururvilla et al. 2024) by examining the date of large hatchery releases and the corresponding estimates of spring-run juveniles at downstream RSTs.

Reconcile hatchery vs natural-sourced salmon in RST efficiency trials

Chapter 4 states that hatchery-origin juveniles tended to have lower capture probabilities even when they were similar in size to natural spring-run juveniles. However, as noted by the MT, the origin of releases was recorded for only 91 of 1,056 trap efficiency trials (9%), and the majority of trials with recorded origin (86) came from a single site (Battle Creek). The database of efficiency trials should be re-examined to make sure natural-origin juveniles were recorded as such when released into rivers where no hatchery releases occurred.

Investigate correlations among tributary and mainstem results, including Red Bluff Diversion Dam (RBDD)

The USFWS began rotary screw trapping at the RBDD site in 1994, making it the longest running mainstem trapping effort in the Sacramento River Basin[3].

---

[3] https://www.fws.gov/project/red-bluff-diversion-dam-juvenile-fish-monitoring

The site is upstream of the independent spring-run populations of Deer Creek, Mill Creek, and Butte Creek, but does capture juvenile production of Clear and Battle creeks and may serve as a suitable surrogate for those creeks if one or the other experiences prolonged outages due to catastrophic events that suspend those tributary monitoring programs. The SRJPE reviewed here doesn't yet utilize this data source.

Evaluate accuracy and precision of spawner counts

Spawner to fry recruitment relationships were often weak and environmental covariates such as temperature and discharge sometimes provided little information. Modeling accounted for uncertainty in the recruitment values but not in the spawner counts. While uncertainty in spawner counts is unknown, the Review Panel recommends using an uncertainty assumption in the model to avoid underestimating it. Uncertainty in spawner counts may vary depending on the method used to estimate them. Factors that might contribute to uncertainty in spring-run spawners include:

1. prespawn mortality while holding in pools prior to spawning,
2. sex ratio, and
3. mixing of fall-run spawners with spring-run spawners.

The GrandTab Escapement Database (https://nrm.dfg.ca.gov/FileHandler.ashx?DocumentID=84381) notes that prespawn mortalities were removed from escapement counts in some tributaries, but the Review Panel was not certain to what extent this occurred in each tributary. It appears the MT assumed 50% of the spawners were female. This should be verified because in some Alaskan watersheds, female Chinook salmon can be less abundant than males, owing to the older age and lower survival of female Chinook salmon (e.g., Olsen et al. 2006). Additionally, redd counts are often multiplied by an assumed factor of 2.5 to account for the greater number of males than females. Lastly, some tributaries include fall-run Chinook salmon that might be potentially counted as spring-run on the spawning grounds. To check on these potential biases, the MT calculated fry per spawner ratios to see if they were reasonable. This evaluation is good, but the Review Panel recommends additional evaluation of the spring-run spawning data.

Develop and publish sampling protocol manuals

Chapter 2 provides an overview of monitoring data availability and refers to another general document about the data (Harvey et al. 2022). If not currently available, detailed manuals should be written for each sampling protocol so that the MT and Review Panel can better understand the data collection methodology and data quality, which are critical for accurate output for the SRJPE. Furthermore, protocol manuals are needed to maintain consistent data collection over time. For example, capture-efficiency trials at each RST are critical for the SRJPE, but the MT noted at the February 24, 2026 *Spring-Run Juvenile Production Estimate Peer Review Public Meeting* that efficiency could range from approximately 25% one week to approximately 2% the next week, without mentioning the typical effect of river flow on efficiency. This high variability raises some immediate questions:

1. Could the method and number of fish released contribute to this high variability?
2. Were marked fish released in a way that they would be randomly mixed with the outmigrating fish by the time they reached the RST?

Question 2a.

Are there additional data/monitoring needs or modifications to existing monitoring approaches that could be considered?

The MT relied upon existing data and monitoring when estimating the SRJPE. They also identified several ways to improve existing data and monitoring. Some of those ideas from the MT are highlighted below, as are ideas from the Review Panel. Please also see comments in the Chapter reviews.

Increase mark/recapture trials at RSTs

Accurate mark/recapture efficiency values are essential for accurate salmon abundance estimates. RST efficiency was approximated by the MT using statistical methods (see Chapter 4/5 comments) when values were not available for a given week. However, the level of uncertainty was very high in these weeks. Ideally, mark/recapture trials should be conducted during each week at each RST, especially during periods when many spring-run juveniles are likely to be migrating (but please see comment about life history strategies in Charge Question 4a). As a cost consideration, trials could be conducted during weeks with high outmigration if staffing is sufficiently flexible. Unfortunately, high outmigration often coincides with high flow when the RST cannot be run (for safety reasons). In these cases, much-improved estimates can be obtained if efficiency trials are conducted immediately before/after large flow events, again requiring staffing flexibility. The Review Panel provides additional suggestions for conducting these trials in the Chapter 4/5 review.

Investigate whether cross-tributary correlations can be used to improve estimates/forecasts

Cross-tributary correlations should be examined for several reasons:

1. High correlations can be used to impute values when data from a tributary are missing for a given year (e.g., due to a forest fire affecting monitoring).
2. High correlations may imply a common environmental factor across all tributaries that could be used as a common covariate.
3. A Bayesian model could be fit to use these cross correlations to reduce the uncertainty in tributary-specific estimates, similar to the effect of “averaging” across tributaries.

Investigate survival and migration rates using small 40-millimeter fish

See the Review Panel discussion of this topic above and in the Chapter 3 and 9 reviews.

Continue/expand genetic testing to support the PLAD model

The presence of numerous fall-run and winter-run in select tributaries can confound length-at-age. Issues with high overlap between spring- and fall-run fish presence and the high uncertainty in early-season estimates could be addressed in part by expanding data collection efforts. Uncertainty in the PLAD models for creeks with more available data decreases, indicating that this is a viable solution for improving estimates on the modeling side if expansions can be made on the monitoring side. The change in stock composition in Battle Creek in 2025, due to the increased presence of winter-run fish, confounded winter- and fall-run classifications, as the model had been trained on three years of data with a different stock composition. This result emphasizes that continued testing at RST sites is necessary to improve PLAD accuracy, as run proportions in a creek can change over time. The review panel also recommends using additional years of data to examine interannual variability and thereby refine PLAD estimates for years that precede the availability of genetic data. However, as future monitoring is considered and implemented, it is important to ensure that the PLAD model continues to use recent data (e.g., data collected in a window of years directly proceeding the development of a PLAD model) so that these changes in stock-composition do not become “baked” into the assumptions for PLAD model results as the run proportions in a creek change overtime due to management decisions, conservation impacts, or environmental influences.

Mark 100% of all hatchery salmon to improve PLAD values

Currently, 100% of hatchery spring-run Chinook salmon are reportedly marked, but only about 25% of fall-run Chinook salmon are marked. Millions of unmarked fall-run hatchery juveniles are released into the Sacramento River basin each year bringing considerable challenge and uncertainty to PLAD-based run assignment for runs that comparatively express only a small fraction of their abundance, like spring-run. Small fall-run hatchery fry are released into the Feather River and can be linked to their parents through Parentage-Based Tagging (PBT). Visually marking hatchery salmon could improve the accuracy of PLAD estimates, which are critical for estimating the SRJPE at the Delta and in spring-run Chinook salmon tributaries where fall-run Chinook salmon are released. Additional comments on this issue are provided below in Charge Question 5.

Coded wire tags (CWT) analysis

Many juvenile Chinook salmon have been tagged with internal CWT tags, largely for fisheries management over many years. We are aware that the MT is currently exploring the potential use of historical CWT recoveries from juveniles captured in the mainstem river and possibly the estuary to estimate travel time from the release hatchery to the estuary. At the February 24, 2026 *Spring-Run Juvenile Production Estimate Peer Review Public Meeting*, the MT noted that most CWT fish were about 70 millimeters, and most were fall- rather than spring-run Chinook salmon. While analysis of these data is worthwhile, it does not fully address the size and timing issues associated with the large acoustically tagged salmon.

Increase the number of RSTs in the mainstem

Capture efficiency of juvenile spring-run Chinook salmon at the mainstem trap sites (approximately 0.25% to 0.47%) is about five to 10 times lower than that in the tributaries (2.5%) because the volume of water flowing by the mainstem traps is much greater. The relationships between capture efficiency and discharge require further examination. For example, what percentage of the river discharge is sampled by the mainstem versus tributary traps, and do capture efficiency rates at each trap make sense given the volume of water sampled by the trap? This type of analysis could be used to determine how many additional traps are needed to more precisely estimate the abundance of the spring-run population at the mainstem sites. Currently, two RSTs are used at each mainstem site, Tisdale and Knights Landing. In contrast, up to five RSTs are used at the Red Bluff diversion dam. In addition to more traps, reasonably accurate estimates from PLAD are needed to identify spring-run versus other run types of Chinook salmon at the mainstem traps.

Evaluate movement into floodplain habitats and its effect on survival and migration timing

Consider evaluating the various flood control structures along the Sacramento River from the Sutter Basin to the Delta entry or the Yolo Bypass and the flows with which they entrain juvenile salmon. At multiple points along the river, structures spill floodwater, and juvenile salmon, into large flood-bypass or floodplain areas that provide growth and possibly survival benefits to the juvenile salmon routed there (Jeffres et al. 2008). Those flood-control features are overtopped at various discrete flow levels. Assessing juvenile migratory speed, growth, or survival at those discrete flow values may yield more information about the relationships of those variables to flow than an assessment based on categorical water year type or peak flows alone. Consider that some of these flood-control structures route fish around the mainstem RST sites currently used in the SRJPE model. Additional monitoring may be required to assess the effects of flood flows on the timing, growth, and survival of spring-run juveniles entering the Delta.

Question 3

Does the proposed SRJPE plan include adequate contingencies for unplanned events such as the loss of adult and juvenile monitoring program data due to extreme events?

The MT documents could benefit from some discussion of the potential impacts of losing future data streams, on both small and large scales, and the effect of those losses on data quality or the integrity of the SRJPE output. There are examples of short-term data loss on the order of days or weeks in the existing historical datasets including harsh weather or flows high enough to make juvenile trapping programs pause for safety, limited spawning ground surveys for storms or fires, reduced trapping effort when programs are occasionally understaffed (shorter trapping seasons or trapping less than 7 days a week), or government shutdowns that require affected agencies to temporarily cease work. Longer-term loss of entire program components is a real risk, given funding uncertainties. Contingency planning or anticipatory adjustments are not fully addressed in the documents.

Modeling Team efforts to reconstruct or simulate missing portions of the data portfolio, such as applying PLAD results to years prior to genetic testing, are examples of adjustments that could prove useful for potential missing components in the future. Correlations among tributaries and Bayesian models could be explored to assess the utility of using neighboring tributary data to estimate juvenile production if an entire tributary monitoring program is shut down due to loss of staff or funding (See response to Charge Questions 2 and 2a). The Mainstem Trap site at Red Bluff Diversion Dam should be explored for its utility in estimating production and timing from upstream tributaries as potential insurance against the loss of one or more of those monitoring programs.

Question 4

Are there additional ways to reduce uncertainties (e.g., changes to modeling or monitoring) in the SRJPE model as it relates to:

Question 4a.

Accounting for the diverse life history strategies exhibited by the Central Valley spring-run population (e.g., How can the SRJPE model incorporate yearlings?).

Incorporating yearling life history

BT-SPAS has programs for two age-classes, so yearlings could be incorporated into the BT-SPAS model if data are available. However, the Modeling Team Draft Report (2025) contained very little data on yearling spring-run Chinook salmon, which might represent about 10% of the juvenile population. At the February 24, 2026 *Spring-Run Juvenile Production Estimate Peer Review Public Meeting*, the MT noted that consistent yearling spring-run datasets appear to be limited. The Review Panel recommends documenting the availability of yearling spring-run data so that the incorporation of yearlings into the model can be assessed. Abundances of yearling spring-run Chinook salmon could increase after restoration of habitat at higher elevations or removal of migration barriers, such as the dam on Battle Creek.

Considering life history diversity

The SRJPE model could account for the considerable diversity in migration timing among spring-run populations. Some populations have early, moderate, and/or late timing components, e.g., Mill and Deer creeks, owing in part to their higher elevation and colder environments (Cordoleani et al. 2021). This diversity helps to spread the risk of migrating too early or too late. Each spring-run population should be adequately sampled at the RST sites to account for diversity in migration timing so that management actions can potentially target each strategy. This might imply that it is important to operate the RSTs over relatively long periods to fully account for diverse life-history strategies. Likewise, while it is important to estimate RST capture efficiencies when numerous juveniles are emigrating, it is also important to conduct efficiency tests during periods when less abundant life histories are emigrating.

Further investigation into successful juvenile life histories, as expressed through returning adults, could be valuable for informing management actions. For example, eye-lens (Bell-Tilcock et al. 2021) or otolith microchemistry analyses of adult salmon to inform the value of floodplain or non-natal rearing could illuminate flow or habitat connectivity covariates to estimate the relative abundance and timing among alternate life-history strategies, such as floodplain versus river rearing and migration. Investigations indicate that the relative success of different life histories can vary with wet versus dry years, suggesting that management actions should focus on all life histories rather than on a single life history (Sturrock et al. 2015, 2019).

Flow paths that circumvent Tisdale and Knights Landing trap sites

When the Sacramento River experiences flows high enough to top certain flood control structures, the river spills into flood bypasses and floodplains that are highly productive for juvenile salmon rearing. Some of these flow paths completely circumvent the mainstem traps at Tisdale and Knights Landing. In very wet years, some portion of the spring-run juvenile production from Clear Creek, Battle Creek, Deer Creek, and Mill Creek, is diverted away from the main channel and experiences faster growth than their cohorts in the mainstem yet are not reflected in the capture at the mainstem traps currently included in the model. Likewise, the Tisdale Weir is just downstream of the Tisdale RST site, and spills water into the Sutter Bypass. The Sutter Bypass rejoins near the confluence of Feather River and Sacramento River downstream of the Knights Landing RST site, so some portion of juvenile production passing the Tisdale RST are routed around the Knights Landing RST when the Tisdale Weir is engaged. The MT mentions a mainstem trap further downstream, Delta Entry, that didn’t yet have enough years of data to include in the model. That site, while downstream of the Feather River confluence, can also miss a significant portion of the SRJPE as it’s downstream of the Freemont Weir which spills floodwater into the Yolo Bypass. All of these mainstem sites suffer from some utility at bypass flows.

Question 4b.

The proposed periodic model updates and whether they provide sufficient opportunities for improvement as new monitoring or science (e.g., Parentage-Based Tagging) become available?

The Modeling Team Draft Report (2025) described its effort as a work in progress, even though it has made considerable progress to date using available data. In each chapter, and especially in Chapter 3, the MT identified and recommended many ways in which to improve the SRJPE. These recommendations and those of the Review Panel, should be implemented, when possible, to improve the SRJPE.

Parentage-based tagging (PBT)

PBT is a genetic approach used to identify and monitor salmon populations by determining the parents of individual salmon. PBT could be used to identify hatchery-origin fall-run salmon, including numerous fry that are similar in size to spring-run juveniles, or progeny of fall/spring-run hybrid salmon spawning in the river. PBT analysis begins with genetic sampling of all spawning salmon, including those in hatcheries. Progeny can be linked to their parents through genetic tissue sampling of juveniles or returning adults. PBT is currently used, for example, to identify numerous fall-run Chinook salmon fry released into the Feather River (see response to Question 5 below). PBT is an exceptional tool for identifying the origin of numerous progeny. Still, it does require tissue sampling and genetic analyses of both parents and progeny, rather than visual identification of a marked salmon or wand detection of CWT salmon.

PBT is a powerful new tool that may lead to improved PLAD estimates and, potentially, to improved selection of broodstock for the Feather River fall-run hatchery, which can mistakenly use natural spring-run adults as fall-run broodstock, thus propagating more hybrid salmon. PBT is currently used in hatcheries that release largely unfed fall-run Chinook salmon fry into the Feather and American rivers, for example. PBT should be further explored and evaluated as a means to improve the SRJPE, especially where the presence of fall-run or winter-run Chinook salmon may confound PLAD estimates of fish run type.

Continuation/expansion of genetic run testing for PLAD

Continuation/ expansion of genetic run testing is needed to reduce uncertainty in the PLAD model predictions. For example, 2025 estimates showed unexpected proportions of winter-run fish in Battle Creek. Updating the PLAD model as recent years of data become available can help determine whether this and future anomalies like it is due to an anomalous year or a new run-time composition spread that has just emerged. Updates to the PLAD model to include years as they become available, and potentially to drop previous years as they become obsolete, can keep PLAD assumptions aligned with the run-time composition in the "current" time period and prevent previous ecological states from being baked into PLAD estimates. The PLAD framework is flexible enough to be updated frequently, and the results from PLAD in the Modeling Team Draft Report (2025) indicate that updating the model and keeping it current as data becomes available will lead to continued improvement in PLAD estimates.

Question 4c.

Any other observed concerns or deficiencies in the modeling approach?

Variance decomposition to partition sources of uncertainty

The MT should conduct a variance decomposition to determine the primary sources of uncertainty in the SRJPE and identify data and monitoring that need improvement. Presently, the greatest source of uncertainty in the SRJPE appears to stem from survival and travel time of salmon as they migrate from tributary RSTs to the Delta entry. This uncertainty does not incorporate potential bias associated with the extrapolation of large acoustically-tagged salmon to salmon fry, which represent the majority of the SRJPE. Further analysis would help identify where the most important improvements could be made.

Management actions are unknown

The MT was tasked to generate the SRJPE and to inform life cycle models without knowing specific management actions that might be used to minimize adverse effects on spring-run juveniles. Knowing the potential management actions that might be undertaken could reduce uncertainty by allowing specific parts of the SRJPE to be targeted. It could also reduce the level of monitoring required for the SRJPE. In 2011, a review of life cycle modeling in the Central Valley was conducted, and the panel concluded:

> *"The most useful models are those developed to address a specific question. The question then guides the decisions and judgments made as to the detail needed, what can be greatly simplified or ignored, and the resolution (time and space scales) needed in the model"* (Rose et al. 2011).

Provide juvenile timing and abundance at the tributary level

The current modeling approach seeks to estimate the combined SRJPE across all sub-populations, as well as the timing of the combined run at the Delta entry. Presumably this is to help management decisions that could affect substantial parts of the combined SRJPE. However, smaller subpopulations may be at risk if they are hidden when aggregated. The Review Panel recommends that the predictions include the abundance and timing of individual tributary populations at Delta entry to avoid adversely affecting smaller sub-populations.

Examine additional statistical approaches for using multiple environmental covariates

In several models, it is not possible to fit many covariates simultaneously due to the limited number of available years of data. In many cases, single covariates were individually fit, and measures of goodness-of-fit (e.g., leave one out) were used to “rank” the covariates. Often, it is difficult to determine whether observed differences in fit are statistical artifacts (e.g., due to outliers), small but biologically irrelevant, or small and biologically relevant. Users will need guidance on choosing covariates.

Given this issue, a common tactic in the face of many covariates is to use multivariate ordination, such as Principal Components Analysis, to investigate collinearity among the covariates and possibly use one or two principal components as composite variables that describe dominant patterns of variation among the original covariates. Multivariate ordination would be beneficial for exploring patterns of variation, but we do not suggest using principal components in the model, since it is challenging to relate them back to the original covariates (i.e., eigenvectors or loadings on the components are vary), which makes ecological interpretations difficult.

If a likelihood framework were used, it would be relatively straightforward to use information-theoretic methods (such as Aikake Information Criteria (AIC)) to combine results from multiple models, where each model is given a weight that accounts for complexity and fit. A weighted average of the individual results could be used rather than picking a single model. This approach would also allow the MT to fit more complex models (e.g., multiple covariates) where the increase in fit (from more covariates) may not overcome the fitting penalty (from more covariates), and so more complex models may not be suitable. Multiple model averaging is much more complicated for Bayesian models (e.g., reversible jump Markov Chain Monte Carlo (MCMC)), but simpler approximations are available via Bayesian Model Averaging (BMA; Hoeting et al. 1999).

Implement future modeling approaches already identified by the MT

The SRJPE is considered a work in progress. Throughout the Modeling Team Draft Report (2025), the MT identified several modeling approaches and data analyses for ongoing and near-future analysis, e.g., explore alternate approaches to estimate capture probability at mainstem RST sites (Chapter 5), improving methodologies to provide more reliable estimates of annual spawner abundance in tributaries (Chapter 7). The Review Panel encourages the MT to follow up on the ideas presented in the Modeling Team Draft Report (2025).

Question 5

Are there any suggestions for how to deal with the challenges described for modeling Feather River natural production (described in the In-season Outmigrant Report), which are caused by more complicated run structure than other tributaries and by large releases from the Feather River Hatchery, including unmarked fry release groups?

Weekly estimates of juvenile spring-run abundances at RST sites on the Feather River were not available in the Modeling Team Draft Report (2025) due to complex run type identification issues. Work has been ongoing to apply the PLAD model to data from the Feather River to identify run type. The MT indicated at the February 24, 2026 *Spring-Run Juvenile Production Estimate Peer Review Public Meeting* that they have made good progress. The Review Panel has not examined new analyses since publication of the 2025 report.

Identification of Chinook salmon run types on the Feather River is more challenging than at other RST sites due to a more complicated run structure and the release of large numbers of hatchery-origin spring-run and fall-run juveniles upstream of most RST sites, approximately 2 million spring-run and approximately 2.3 million fall-run juveniles released per year. The run structure of juvenile Chinook salmon is complicated by the presence of natural and hatchery spring-run Chinook salmon, natural and hatchery fall-run Chinook salmon, and spring/fall-run hybrids produced by adults interbreeding on the spawning grounds and by unintentional inclusion of untagged spring-run Chinook salmon in the fall-run hatchery broodstock. 100% of hatchery spring-run Chinook salmon are typically marked, sometimes only 50%, but only 25% of hatchery fall-run Chinook salmon are marked. Many hatchery fall-run Chinook salmon are now released as unmarked fry, which must be identified by genetic testing, Parentage-Based Tagging, to approximate spring- versus fall-run salmon in the migration.

According to the Modeling Team Draft Report (2025), the policy for outmigrant run assignment in the Feather River is to use run-time specific genetic markers to assign fish homozygous for early-running genotype to spring-run and those homozygous for late-running genotype to fall-run. Juveniles heterozygous for the run-timing markers are assigned to either spring- or fall-run using a suite of other genetic markers, which are not associated with run-timing but are based on similarities or differences in genetic composition with the spring-run population or fall-run population. These heterozygous early/late juveniles may emigrate across a wide range of sizes and dates. As a result, individuals assigned to spring-run by the population-based assignment method may emigrate at sizes and dates more typical of fall-run, and vice versa. This poses a problem for fitting PLAD models, which are trained on recent years of data with genetic outmigrant assignments and then used to assign run type to catch data from years prior to genetic testing. Given that most years of data predate genetic testing, PLAD-based assignments could introduce substantial bias in spring-run outmigrant abundance and timing estimates for those years. Furthermore, a major limitation in modeling stock-recruit relationships in the Feather River is the lack of naturally spawning spring-run adult abundance estimates before 2014 to match RST data from 1998–2013. Given that the Feather River is either the largest or second-largest producer of naturally-spawned spring-run Chinook salmon in the Sacramento River watershed, these issues could have a substantial influence on the accuracy of the valley-wide SRJPE. Considering these complex issues in identifying Chinook salmon run types, the Review Panel has several recommendations in the paragraphs below.

Separate the Feather River spring-run juvenile production estimate from the valley-wide SRJPE. The naturally spawned spring-run Chinook salmon population in the Feather River is relatively large, and it will likely face greater uncertainty than other populations due to the variety of complex run-type identification issues noted above. Including Feather River in the valley-wide SRJPE would increase uncertainty in the SRJPE, especially for weekly abundances entering the Delta. Furthermore, while newly proposed methods might reduce uncertainty, see below, the Review Panel suspects that uncertainty of the Feather River estimate will remain high relative to other spring-run populations.

Timing of potential management actions intended to benefit spring-run Chinook salmon could be misinformed by the inclusion of Feather River salmon, which strongly influences the overall SRJPE and its seasonal timing. For example, management actions might target the period when juvenile abundance entering the Delta is high, due to Feather River salmon, even though peak migrations of less-abundant populations may occur at other times. Rather than focusing on a valley-wide SRJPE, the Review Panel recommends that managers account for the diverse timing of spring-run populations, including early- and late-migration components, when considering actions that might benefit spring-run juvenile Chinook salmon. Protecting population diversity is an important aspect of maintaining salmon viability, especially in a changing environment.

The MT also identified several new approaches to estimate Feather River spring-run juvenile Chinook salmon abundances and timing potentially. Some of these are listed below, while others can be reviewed in Section 5 of the MT Chapter 3. The Review Panel supports these new approaches.

Model heterozygotes explicitly

The MT expects that heterozygous fish will exhibit intermediate outmigration length-at-date characteristics relative to fish with homozygous early and late genotypes. This may allow estimation of the proportion of heterozygous fish in RST data collected before regular genetic testing, which may help improve the accuracy of PLAD models. This approach would still require some means for assigning heterozygous fish in historical catch data to either spring-run or fall-run, if estimates are extended to years before genetic testing. The Review

Panel notes that the uncertainty of these estimates must be evaluated to ensure it does not overwhelm values in the valley-wide SRJPE.

Use Eye Riffle RST data instead of Herringer Riffle

The MT proposes switching BT-SPAS-X modeling from Herringer to Eye Riffle, plus nearby Steep Riffle and Gateway Riffle, to align with the location of genetic testing. The proposed RST sites are in the low-flow channel downstream of spring-run spawning locations, and they are close enough together that the PLAD model based on genetic sampling at Eye Riffle is expected to apply to all three sites. Importantly, the Eye Riffle RST complex is upstream of the site where hatchery salmon are released. Thus, species identification only involves spring-run, fall-run, and hybrids that spawn in the low-flow channel upstream of Eye Riffle. Hatchery production of spring-run salmon would be estimated from the number released and estimated survival during emigration.

The Review Panel generally supports this approach for separating natural production from hatchery production, although more details would be beneficial. This approach would reduce the number of years of available data, but additional years of data are likely forthcoming. Importantly, the Eye Riffle spring-run estimates would likely be much more accurate than those stemming from the Herringer RST where numerous spring- and fall-run hatchery salmon complicate the estimates. Hatchery production of both run types is exceptional, and the release of numerous unmarked fall-run hatchery salmon fry requires parentage-based-tagging to estimate spring-run juveniles at the Herringer site. Nevertheless, the MT still plans to provide abundance estimates from the Herringer site, apparently, to maintain the time series back in time.

An alternative and potentially more accurate approach to identifying hatchery salmon at the Herringer site is to mass mark 100% of fall-run Chinook salmon, including pre-smolt and smolts, along with 100% of spring-run hatchery salmon, so they can be readily identified in the RST catch. Also, 100% marking of hatchery salmon could eliminate the production of hybrid Chinook salmon, which occurs when unmarked spring-run Chinook salmon are unintentionally used as broodstock along with unmarked and marked fall-run Chinook salmon. However, it appears that the hatchery prefers to identify fall Chinook salmon fry using PBT and to mark only 25% of the larger fall-run Chinook salmon, an approach that requires greater effort when identifying spring-run salmon.

Public Comment

Two members of the public spoke at the February 24, 2026 *Spring-Run Juvenile Production Estimate Peer Review Public Meeting* in Sacramento. Both urged the Review Panel and SRJPE team to consider the big picture and the importance of this effort to spring-run. Are the questions posed to the SRJPE team framed in the best way possible to manage a water system that serves millions of people while protecting spring-run? Specific management actions for evaluation should be defined. Understanding how this information will be applied is important.

The Review Panel agrees with the public input. The Panel recognizes that the JPE Core and Modeling Teams were directed to develop the SRJPE tools without knowing what form of life-cycle model and/or minimization measure would take. This is a significant disadvantage faced by the Model Team. Additional refinement and improvement of the SRJPE could be made when this information becomes available.

Review of Material and Supplemental Documents

Common concerns about all modeling chapters

The MT has a difficult task of summarizing and explaining very technical statistical models (e.g., hierarchical models) with complex fitting procedures (e.g., MCMC methods in a Bayesian context) used in the submodels. Consequently, each chapter must be written carefully and concisely so that the methods used can be understood and reproduced. The Review Panel has identified several issues, common to many chapters, that could impede readers' understanding of the chapters. Additionally, the Review Panel had many questions for the MT to consider as they improve and advance the SRJPE. Those questions are scattered through the individual chapter reviews here and in Appendix A.

Statistical models require careful attention to notation. The Review Panel found that notation was often confusing. For example, some of the technical descriptions of the models used proper mathematical notation, some used shorthand notations, and some used computer code notation, etc. Each chapter should be carefully edited to use standard mathematical notation with proper subscript/superscripting and distribution names.

The MT needs to handle large volumes of data and often consider multiple models for the same data, e.g., by investigating different covariates. The chapters lack basic data summaries, preliminary assessments of the data (e.g., plots and tables), and preliminary relationships (e.g., plots/tables of raw capture efficiencies versus flow) that summarize the data and justify modeling choices.

The submodels should use a consistent, standardized date system so that date-specific information from one submodel is correctly linked to the date in another model. For example, the PLAD model in Chapter 6 uses days since July first to predict Chinook salmon run type, but this start date precedes the juvenile spring-run start date. Also, the outmigrant forecast model in Chapter 8 defines the migration season as November 1 in year $t-1$ to October 31 in year t. These dates differ from the water years used by hydrologists. The current MT likely has dates aligned correctly, but 10 years from now, model implementers may not know these basic alignments.

The Bayesian models that are fit are quite complex. Documentation for any model requires a full description and justification of the prior distributions used for parameters, a complete description of the likelihood, a full specification and justification of random effects often used in hierarchical models, a complete description and justification of derived parameters (e.g., total abundance of a run as the sum of weekly abundances), and a description of how uncertainty is propagated.

Similarly, after the models are fit, the MT appears to rely on a single measure of model assessment (Rhat; Gelman and Rubin 1992; Brooks and Gelman 1998). This statistic assesses convergence by comparing within-chain variance to between-chain variance of a parameter. Unfortunately, Rhat could indicate that a model has “converged”, but the model may be a poor fit to the data, analogous to a linear fit to data when there is evident curvature. There are many other measures of model assessment that could also be examined that assess other aspects of the fit:

1. Traceplots showing mixing and autocorrelation of parameters across successive iterations.
2. Autocorrelation plots (ACF) which may indicate that more severe thinning is needed and/or a reparameterization to avoid high correlation among parameters.
3. Prior-posterior overlap (PPO). This is an assessment of the degree to which the prior is informing the posterior distribution. Substantial influence of the prior on the posterior may not be apparent through the use of R-hat and visual checks alone. Checking the PPO is particularly useful for determining whether the model's parameters are identifiable. If substantial PPO exists, the prior may simply be dictating the posterior distribution, and the data may have little influence on the results. Gimenez et al. (2009) developed quantitative guidelines for determining when parameters are robustly identifiable based on prior-posterior overlap.

4. Posterior predictive plots (PPP), or Bayesian p-values, looking at various aspects of fit. Given a model M, the PPP examines how well it fits the observed data $xobs$ based on a statistic T, which measures the goodness of fit of data and model. For example, the tributary BT-SPAS-X models fit a hierarchical model over all years and weeks, but did not include a year random effect. How do simulated values from this model compare to the actual observed efficiency trials? An example of computing a posterior predictive plot based on efficiency trials is found in Appendix B.
5. Plots of the mean of posteriors versus actual data. For example, BT-SPAS-X fits a spline curve to the run, but there do not appear to be plots of the actual spline fit on the logarithmic scale. This could help identify, for example, weeks with a large jump in the run, such as after a rain event, indicating that a model with a discontinuity in the run curve may be needed. Similarly, the plot of the fitted spline may indicate that the sampling for a run started too late or ended too early.
6. The MT approximated a posterior distribution for a parameter in one stage, for example, using a normal distribution, to serve as a “prior” for the parameter in the next stage of the multi-stage model. This needs justification, e.g., plots of the posterior with a normal distribution overlaid. In particular, does the normal approximation have sufficiently fat tails to capture the true uncertainty expressed by the posterior?
7. The MT made choices in the model structure, cubic versus quadratic splines or second-order random walks. These choices need justification.
8. The MT assumed that all submodels are independent. This does not appear to be assessed. If a correlation, e.g., between abundance and survival was found, this may indicate a “higher” order covariate that should be considered for both stages, e.g., cooler/wetter years both lead to better survival in rearing and survival downstream.

To avoid creating a document that is too large, consider creating two reports for each chapter. One report should be written as a scientific publication where a concise description is given of the model with the results obtained, and additional reports should be written as “web appendix” where preliminary plots, model assessment, and other “working” information should be presented for reference if needed.

The documents appear to have been prepared using Microsoft Word, or similar word processing software. These reports are dynamic and will be updated each year as more data are collected. This is a time-consuming and error-prone process. Consequently, the MT should consider using Quarto (https://quarto.org/; an open-source technical publishing system) that allows writers to create dynamic documents that “automatically” update themselves as new data are obtained. It also allows the creation of documentation in a variety of formats, such as PDF or HTML, with very little extra work.

Chapter 1 Introduction and IPR Materials Package Overview

Chapter Overview

Chapter 1 offers a broad overview of the scientific framework, objectives, and modeling approach developed to estimate juvenile production of spring-run Chinook salmon entering the Sacramento–San Joaquin Delta. Guided by a structured decision-making process and interagency collaboration, the chapter describes a series of clearly articulated management objectives that shaped the development of a flexible, integrated modeling framework. In brief, the objectives emphasized developing a juvenile production estimate that is reliable, timely, and useful for management, while remaining feasible to implement. The stated objectives prioritized maximizing confidence in estimates of juvenile abundance and timing at Delta entry, while accounting for life-history diversity and inclusion of all spring-run tributary populations. Additionally, the objectives sought to balance scientific rigor with practical considerations by minimizing impacts on listed species, controlling costs, and ensuring that monitoring could be conducted consistently as planned.

Chapter 1 recognizes that data availability and forecast timing are key considerations for developing a SRJPE, leaving it to subsequent chapters to explain how differences in data availability would influence confidence in the forecast. This reflects the practical constraint, as articulated by the MT, that the form of future minimization measures could not be anticipated. Therefore, a single forecast timing requirement could not be defined. Rather than limiting applicability, the chapter describes how this uncertainty motivated the selection of a flexible, integrated modeling framework. By emphasizing adaptability and the ability to incorporate all available data at any given forecast date, the Review Panel believes that the integrated modeling approach remains well positioned to support a range of future management applications in light of continued evolution of data needs and decision contexts. Nevertheless, the Review Panel notes that model improvements could be made once potential management actions are known.

Recommendations

The Review Panel concludes that Chapter 1 sufficiently outlines the major steps of the SRJPE modeling process and explains how the individual submodels fit together within the integrated approach. It offers a high-level summary of the decision context, data availability, and tradeoffs considered, providing the necessary context and transparency to support independent peer review of the proposed modeling framework. No recommendations are provided.

Chapter 2 Monitoring and Data Overview

Chapter Overview

Chapter 2 provides an overview of available monitoring data and the process for updating model datasets to forecast the SRJPE. Monitoring data are summarized, and the procedures for updating datasets are described to provide a clear framework for modeling. Data collection sites, the types of data collected, and program leads are presented in accompanying figures and tables, offering an organized overview of the monitoring network. Compiled and processed datasets are publicly accessible through the *SRJPEdata* R package on GitHub, ensuring transparency and ease of use for modeling and analysis.

Juvenile abundance is estimated primarily using rotary screw trap (RST) data, with information from multiple tributary sites combined in the BT-SPAS-X models to reduce uncertainty in model outputs for individual locations. Newly established mainstem RST sites, including those on the lower Sacramento and Feather rivers, are not yet included in the modeling due to limited data availability. Adult monitoring data, such as video passage counts, holding counts, carcass estimates, and redd counts, are used to support stock-recruit models, while physical covariates, including flow and temperature measurements, serve as important predictors across various modeling frameworks. The curation and preparation of these data follow a structured five-phase process:

1. discovery and outreach to identify datasets and capture metadata,
2. quality control and exploration to assess data integrity and consistency,
3. publication of the data to the Environmental Data Initiative for public access,
4. standardization and integration into the *SRJPEdata* package to enable seamless use across models, and
5. detailed data processing, including species filtering, removal of adults and yearlings, application of length-at-date rules, and aggregation to weekly timesteps. Covariates to support both historical analyses and forecast models were well described.

The chapter also provides details on the process by which datasets are updated and maintained. Specifically, DataTackle, a tablet-based application, is described as enabling near-real-time upload of field data to a cloud database. These data can be integrated into *SRJPEdata* to support modeling, with updates occurring biweekly for juvenile RST data, annually for adult data, and weekly during the outmigration season to support in-season forecasts. The goal, as articulated by the authors, is for the SRJPE models to be adaptive, readily incorporating newly collected data as they become available. This approach appears to allow models to be updated regularly, from refitting coefficients with updated datasets based on best available science to testing new model structures and covariates, ensuring that forecasts reflect the most current understanding of the Chinook salmon ecosystems.

Recommendations

The Review Panel concludes that Chapter 2, which reports all data collected and processed for the SRJPE, was documented, but details of the sampling methodologies were not discussed or cited.

If not currently available, detailed manuals should be written for each sampling protocol so that the MT and Review Panel can better understand the data collection methodology and data quality, which are critical for accurate and precise output for the SRJPE. Furthermore, protocol manuals are needed to maintain consistent data collection over time.

A notable mainstem trapping location not included in the SRJPE model is the USFWS site at Red Bluff Diversion Dam (RBDD). It's the longest operating RST site in the Sacramento River basin, and as many as five RSTs are utilized to achieve much higher capture probabilities than the Tisdale or Knights Landing sites. The Tisdale and Knights Landing sites both miss portions of the recruitment from upstream when flood flows route some fish around those trap sites. The Mainstem site at Red Bluff Diversion Dam does not face that same issue, there is no flood bypass around that trap location. While Red Bluff may be farther upstream than ideal for quantifying the total spring-run entering the Delta, it could still be useful as insurance against the loss of one or more upstream tributary monitoring programs.

Chapter 3: Integrated Spring-run Chinook Salmon Juvenile Production Model

Chapter Overview

Chapter 3 provides an overview and demonstration of the SRJPE framework. It integrates predictions from three submodels:

1. a stock-recruit model that provides an early/preliminary annual forecast of annual outmigration at rotary screw traps (RSTs) based on spawner abundances and physical covariates (flow, water temperature, or discretized versions of these; see Chapter 7),
2. an in-season model using cumulative juvenile migration data over time to estimate outmigration abundance at each RST (see Chapter 8), and

3. a Cormack-Jolly-Seber (CJS) juvenile survival and travel time model to forecast survival rates from RST locations to Delta entry (see Chapter 9).

These submodels used results from additional submodels that estimated total juvenile salmon abundances at tributary and mainstem sites using RST catch data, Chapters 4 and 5, and PLAD estimates, Chapter 6, to assign run identification to estimate spring-run juvenile abundance.

Historical data were used to demonstrate model behavior, showing that stock-recruit predictions provided useful prior information early in the season. At the same time, in-season estimates reduced uncertainty later in migration, particularly for Butte Creek, which dominated total abundance estimates because Feather River and Yuba River were not included in the SRJPE due to run-identification complications. Chapter 3 also describes how the SRJPE provides pre-season predictions of juvenile arrival timing in the Delta for each tributary population, an important model output that can inform managers about the arrival of less abundant, potentially more vulnerable populations.

According to the MT, Chapter 3 demonstrates how the SRJPE process works, but is not yet ready to provide in-season juvenile spring-run abundance forecasts. Estimates of juvenile spring-run outmigration abundance at the RST sites and at Delta entry should be considered preliminary owing to several issues, including bias in Yuba juvenile estimates and the lack of Feather River data.

General Comments on SRJPE

The MT conducted innovative and well-conceived analyses that used much of the available data. However, as recognized by the MT and discussed in each submodel chapter and in the Review Panel, there were several sources of bias and uncertainty of the SRJPE estimates. At the RST sites, mark-recapture trials were limited in some tributaries and weeks, and it appears traps may not have been operating on some days within some weeks, the number of days when fish were not counted in the traps was not clearly reported. BT-SPAS-X attempted to account for missing values, but these estimates led to greater uncertainty. Additionally, survival and migration timing from the tributaries to the Delta entry relied on acoustically-tagged smolts (spring-run and fall-run) whose median size (89 millimeters) and minimum size (approximately 70 millimeters) were much greater than the size of most spring-run Chinook salmon (approximately 37 to 40 millimeters), which also emigrated much earlier than the tagged fish. The model attempted to account for fish size and flow effects by extrapolating to fish sizes much smaller than those used in empirical studies. Thus, the bias and uncertainty of abundance estimates and timing estimates at Delta entry may be larger than indicated by the SRJPE model. This topic, including recommendations to address this issue, is further discussed in the Charge Questions and in Chapter 9.

Modeling Comments Involving the SRJPE

Propagation of uncertainty is difficult when integrating all the submodels that contribute to the SRJPE. Each submodel uses a Bayesian model to analyze data from multiple years. This yields a sample from the posterior distribution of parameters of interest (e.g., mean travel time) and posterior predictive distributions of individual-based values (e.g., actual travel time for a fish) for a new forecast year. How can these posterior distributions and posterior predictive distributions be combined?

The authors often use a normal distribution to approximate a posterior/posterior predictive distribution rather than using samples from the distributions directly. This may be reasonable, but it needs justification, presumably in the submodel chapters. For example, the shape of the posterior could be decidedly non-normal early in the season when RST data are sparse. Chapter 3 Figure 9 (Figure 1 in this report) illustrates the problem – the posterior distribution of survival from RST to Delta entry clearly does not have a normal distribution.

By using an approximation for the posteriors, the combined model is relatively simple to implement. Indeed, a single page of code, Appendix A of the MT Chapter 3 report, represents the basic computation engine.

The alternative to using an approximate normal distribution is to use the samples from the posterior and posterior predictive distributions directly. Combining posterior distributions is known as a Bayesian two-stage estimation process and is well described in the literature. This can be tricky, but several discussions on the Stan forum will be helpful. Given the large uncertainty in the final SRJPE results, moving away from the approximate normal distribution may not be advisable at this time.

While the modeling for survival to Delta entry is fairly clear, the model structure for arrival timing is much less clear. It requires some careful editing and additional text to help the reader understand this part of the model.

There is some confusion about how variability/uncertainty in travel time is modeled in this chapter. In Chapter 9, a model was built to predict travel time (TT) from the RST to the Delta entry as a function of fish size and flow. Furthermore, the model in Chapter 9 provides a measure of fish-to-fish variability in the travel time. This leads to a posterior distribution for the parameters of the travel time model and posterior predictive distribution for the travel times of individual fish.

However, Chapter 3 Equation 10 says that TT is a posterior distribution rather than a posterior predictive distribution. Which parameter is described by the posterior distribution for TT? Chapter 3 Figures 11/12 shows shaded credible intervals, presumably for individual fish? Is this the distribution of TT that is referred to here?

Chapter 3 Equation 10 appears to model individual fish travel times, but Chapter 3 Equation 11 models the entire abundance in the week using *pD* and the survival probability for that week, with no reference to the TT distribution. It is not clear how the variability in TT is propagated here. Presumably, the variability occurs in the arrival week at the Delta entry for the smolt (*arr_wk* in Chapter 3 Equation 10)?

Because fish spread out as they move, the arrival week consists of fish from several departure weeks. Presumably, the “array” that is alluded to in Chapter 3 Equation 11 gathers the individual fish. A simple example showing the “array” that is alluded to and how it is populated would be helpful here. Chapter 3 Appendix A, source code, also does not appear to reference the travel time computations.

The distinction between a posterior distribution and a posterior predictive distribution is not a problem when modeling survival to Delta entry. The law of large numbers implies that the distribution of survivors can be approximated by the initial number alive multiplied by the survival probability, so individual fish need not be tracked.

In the previous section, Chapter 3 Section 2.1.2, Chapter 3 Equation 7, normal distributions were used to approximate the posterior or posterior predictive distributions. In the section on travel times, there is no mention of this. Does this mean that the posterior predictive distribution for travel times was used directly from Chapter 9?

This chapter does not use the information from the mainstem RST for the SRJPE. This needs to be explained. Presumably, the very low efficiency in the mainstem RST implies that the estimates have very high uncertainty and may not provide useful information.

There is no information provided on how the submodel results are stored. For example, is the complete sample from a posterior or posterior predictive distribution stored, or are only summary statistics stored? How will the actual model and its output distribution be kept in sync over time?

There is an implicit assumption of independence among submodels. This needs justification. For example, perhaps higher abundance is associated with higher survival to Delta entry due to common environmental conditions such as flow and temperature.

A discussion of the fragility and resiliency of the integrated model is needed. For example, what happens if a RST is lost for the year? Presumably, some RST datasets on larger tributaries are more important for total abundance calculations. Cross-tributary correlations need to be examined for potential data imputations. For example, if the Butte Creek population has an estimate but other RST sites are missing an estimate, could it be used as a proxy or stand in for sites or years where there are data gaps? In Chapter 3 Table 2, it is clear that Butte Creek is skewing the aggregate time of Delta entry estimates, owing to its relatively abundant population. Can Butte Creek or other RST sites be used for redundancy against potential anomalies, disasters, or funding cuts at other sites? Given the independent status of Deer, Mill and Butte creek spring-run populations, the Review Panel recommends that these populations remain prioritized.

What will be the potential impacts of climate change on the monitoring program? For example, flashy systems with large rainfall events may make monitoring with RST more difficult. Cross-tributary correlations could be beneficial when a specific RST is not operating.

The final estimate of abundance at Delta entry is very uncertain, and this uncertainty does not incorporate the assumption that large acoustically-tagged salmon represent the majority of the spring-run emigration, which involves smaller fish that migrate earlier than the tagged fish (see below). The report should examine components of variance in the submodels to know where to concentrate future effort. For example, if 90% of uncertainty comes from estimating survival from RST to Delta entry, then this needs further work.

Feather River and Yuba River Estimates

Estimates of juvenile spring-run abundances at RST sites on the Feather River were not available in the Modeling Team Draft Report (2025) due to complex run identification issues. Estimates of Yuba River juveniles were provided, but the MT concluded that the estimates exhibited substantial positive bias, so they did not include Yuba River juveniles in the SRJPE.

Work by the MT has been ongoing to apply the PLAD model to data from the Feather River to identify run type. Likewise, the MT is re-examining run identification predictions from the PLAD model for the Yuba River. The MT indicated at the February 24, 2026 *Spring-Run Juvenile Production Estimate Peer Review Public Meeting* that they have made good progress on these issues. The Review Panel has not examined the new analyses since publication of the 2025 report.

The Review Panel provides more information on the Feather River in its response to Charge Question 5.

Recommendations

In summary, SRJPE can approximate the true number of spring-run juveniles entering the Delta within its current limitations, providing a useful preliminary estimate with associated uncertainty. The framework provides credible intervals to quantify uncertainty, but these intervals do not fully capture all sources of bias or uncertainty from missing data. Out-of-sample testing, model refinement, verification of acoustically tagged salmon, inclusion of additional tributary data, and improved covariate forecasting are needed to better align juvenile production estimates with true annual juvenile abundance and to more accurately quantify the uncertainty around these estimates.

Continue to provide estimates for each spring-run population

Rather than focusing on a single valley-wide SRJPE, the Review Panel recommends that managers account for the diverse timing of spring-run populations, including early- and late-season components, when considering actions that might benefit spring-run juvenile Chinook salmon. Chapter 3 Table 2 shows that spring-run populations have median dates of Delta entry ranging from December 21 (Clear Creek) to February 8 (Butte Creek) to April 5 (Deer Creek). Furthermore, some independent populations (Deer and Mill creeks) produce juveniles that emigrate over a relatively long period of time (Cordoleani et al. 2021). Protecting population diversity is an important aspect of maintaining salmon viability, especially in a changing climate.

Enhance certainty in survival rates from RSTs to the Delta

A significant source of uncertainty in SRJPE estimates arises from predicting juvenile survival between tributary RST sites and Delta entry. The MT needs to verify whether the large acoustically-tagged salmon (median size 89 millimeters) accurately represent the much smaller and earlier-migrating juveniles (approximately 37-40 millimeters) that are the primary component of the spring-run populations. See Charge Question 2a for suggestions on how this might be accomplished.

Explore potential biases in mainstem RST catch data

The reported SRJPE predictions highlight notable discrepancies between tributary-based forecasts and observed passage timing at mainstem RST sites like Tisdale and Knights Landing. These discrepancies are likely due in part to low and variable trap efficiencies, which reduce confidence in mainstem catch-based estimates. As noted in the Charge Question section, the managers should consider increasing the number of RST at Tisdale and Knights Landing to increase trap efficiency in the large river. For example, up to five RSTs are deployed at the Red Bluff Diversion Dam.

Enhance covariate forecasts for environmental conditions

SRJPE relies on covariates such as water temperature, flow, and other environmental factors to reduce bias and uncertainty of abundance and survival estimates. Currently, forecasts use historical covariate data or annual averages, which may not capture the week-to-week variation that influences juvenile outmigration timing and survival. In many cases, it may be possible to discretize environmental factors (e.g., low, moderate, high flows or temperature bands) to generate scenario-based predictions; however, more continuous hydrographs simulated from discharge data would be optimal. The MT recognizes that this issue needs further evaluation.

The Review Panel provides information on the effects of higher flows on the survival of juvenile fall-run Chinook salmon later in this document. High flows may speed salmon travel rates in the mainstem and reduce exposure to predators, or high flows may allow juveniles to enter floodplain habitats where they can feed and grow rapidly. Sacramento River flows of about 38,000 cubic feet per second at Tisdale spill floodwaters to the Sutter Bypass via the Tisdale Weir, for example. Multiple other flood control structures spill at discrete mainstem flows, exposing emigrating spring-run juveniles to significant increases in rearing habitat with each connection to an off-channel bypass or floodplain. The relationship of survival and migration timing to those discrete flow values should be explored with estimates derived from fish of appropriate size to represent the sources from which they come (e.g., Clear Creek fish are mostly smaller than can be acoustically tagged).

Figure 2 of this report is based on unpublished California Department of Fish and Wildlife (CDFW) and USFWS fall-run Chinook salmon data. The Review Panel put the figures together, exploring the relationships between Sacramento Water Year Index and cohort replacement rate, and Water Year Index and the number of adults that returned to Clear Creek three years later. Most Sacramento River Chinook salmon adults return at age three. Although these relationships do not estimate juvenile survival from Clear Creek to the Delta, they do indicate higher survival in wet years. There are far less data of this type for spring-run Chinook salmon, but the Review Panel expects similar relationships.

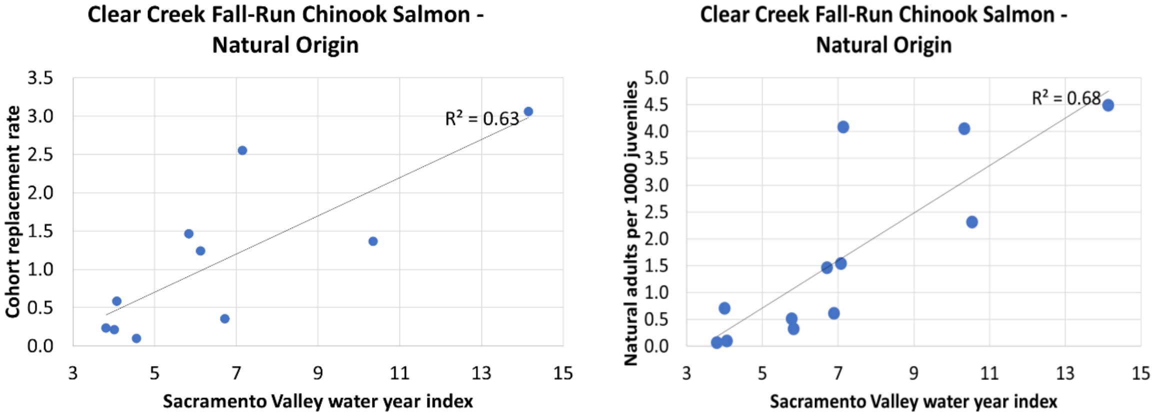


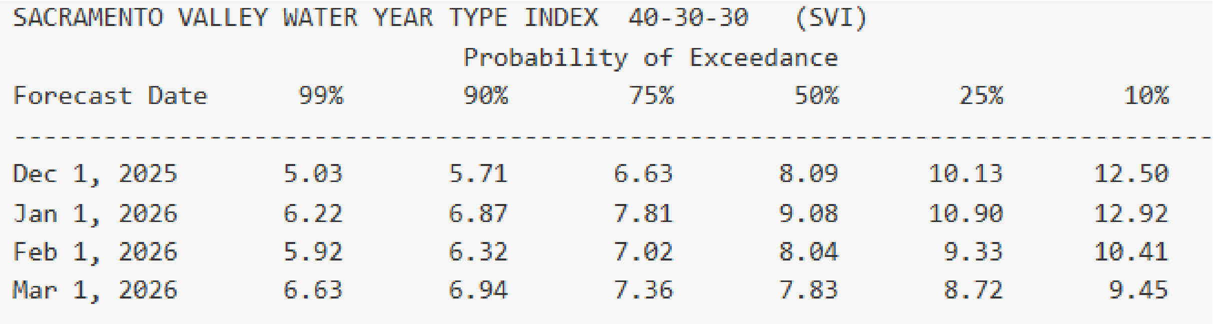
Figure 2. Relationships of adult natural-origin fall-run Chinook salmon entering Clear Creek, to the Sacramento Valley water year index those fish experienced as juveniles while they reared and migrated downstream (assuming age 3 adults). Cohort replacement rate on the left, and adults per estimated 1,000 juveniles produced. Unpublished natural-origin adult salmon data from CDFW, and unpublished juvenile passage data from USFWS.

The mechanistic relationship between flow and survival is often thought of as one in which high flows swiftly route juveniles to the Delta or ocean, reducing mortality on their temporally shortened journey. In high flow years, a significant portion of the population can indeed transport out of the river quickly, however, the mechanisms of survival are likely far more complicated than that. High flows also spill onto highly productive floodplains and bypass reaches, where emigration is slowed but food resources are high, growth rates increase, and exposure to predators such as Striped Bass is likely reduced. Size plays a significant role in survival at sea (Woodson et al. 2013). There is a risk/reward relationship that occurs on floodplains too, as some fish may be lost to stranding or other mortality. However, a high proportion of returning adults exhibit a floodplain rearing history, as reconstructed from otolith or eye lens chemistry analyses (Bell-Tilcock et al. 2021).

Forecasting peak flow may be less valuable than forecasting floodplain connectivity, bypass connectivity, or the utilization of non-natal tributaries. Phillis et al. (2018) observed that 45 to 65% of returning adult winter-run Chinook salmon expressed rearing in non-natal habitats downstream of their natal origin, including other tributaries or off-channel habitats. Off-channel habitats can provide greater Chinook salmon growth through temperature and/or prey abundance benefits (Limm and Marchetti 2009). Connectivity to these habitats is greater in wetter years. There may be a stronger relationship between the continuous variable of Sacramento River Water Year index forecasting and fish survival to adulthood, than there is the categorical carriable of forecasted water year type and fish survival to adulthood.

Water Year runs from October 1 to September 30 each year. The wettest months in the Sacramento Basin are December through April. Beginning in December each year and running through May, the DWR provides monthly forecasts for Sacramento Valley and San Joaquin Water Year Index (https://cdec.water.ca.gov/reportapp/javareports?name=WSI). The Index forecast is reported in 99%, 90%, 50%, 25%, and 10% exceedances, providing a means to include error (Table 1). High and low probability of exceedance forecasts (e.g., 90% and 10%) tend to be widely dispersed from the 50% exceedance forecast at the beginning of the series when few of the typically wet months of the year have been encountered, and then converge closer to the 50% forecast as additional months of the wet portion of the year are accrued until the final forecast is delivered in May. Uncertainty tends to narrow with the addition of each month.

Table 1. Monthly Sacramento Valley Water Year Type Index forecasts through March of water year 2026 from https://cdec.water.ca.gov/reportapp/javareports?name=WSI.



```
SACRAMENTO VALLEY WATER YEAR TYPE INDEX  40-30-30   (SVI)
                              Probability of Exceedance
Forecast Date       99%        90%        75%        50%        25%        10%
--------------------------------------------------------------------------------
Dec 1, 2025        5.03       5.71       6.63       8.09      10.13      12.50
Jan 1, 2026        6.22       6.87       7.81       9.08      10.90      12.92
Feb 1, 2026        5.92       6.32       7.02       8.04       9.33      10.41
Mar 1, 2026        6.63       6.94       7.36       7.83       8.72       9.45
```

Chapter 4 BT-SPAS-X Tributary Outmigrant Abundance Model at Rotary Screw Traps, and Chapter 5 BT-SPAS-X Mainstem Outmigrant Abundance Model at Rotary Screw Traps

Chapter Overview

These chapters describe BT-SPAS-X, an enhanced model designed to estimate both weekly and annual abundance of outmigrating Chinook salmon juveniles, including totals for all Chinook salmon run types combined as well as run-specific estimates for spring-run, using RST data from Sacramento River tributaries and two mainstem sites.

BT-SPAS-X uses a multi-step approach:

1. A Bayesian hierarchical model is fit to the efficiency trials. Several covariates are examined to help explain variation in efficiency across weeks, years, and sites.
2. A Bayesian penalized spline model is fit to estimate the weekly abundance of juveniles at each RST site using the number of unmarked fish caught each week, expanded by the imputed capture-efficiency for weeks where the efficiency trials were conducted and for weeks when efficiency trials were not conducted. The spline model is also used to interpolate for weeks when the RST was not operating. This provides an estimate of outmigrants of all run types.
3. Finally, this outmigrant estimate (all run types) is scaled by the proportion of spring-run juveniles among all run types, which is computed by the PLAD model (reviewed in chapter 06) to give a spring-run only weekly abundance.
4. This final weekly abundance is summed over all weeks to provide an estimate of the SRJP at the RST.

The main difference between BT-SPAS-X and the previously published BT-SPAS model (Bonner and Schwarz 2011) is that instead of using each year’s data separately, the BT-SPAS-X model uses a hierarchical model for capture efficiency across all years and sites. This model is fit separately, and the posterior distributions are used as “priors” in the spline model, which was similar to that used in BT-SPAS, with a minor change in the maximum run size allowed in each week. Hence rather than a single combined model, BT-SPAS, providing estimates of capture efficiency and abundance, two sequential stages are used for BT-SPAS-X.

BT-SPAS-X was applied to 15 RST sites across seven Sacramento River tributaries and two mainstem sites, Knights Landing, Tisdale, to estimate weekly and annual outmigrant abundance of all Chinook salmon run types.

We recommend combining the two chapters into a single chapter because the modeling framework is common to both.

Preliminary data processing

There was very little description of the RST protocols across the years and sites. For example, do the RST programs follow a single-site RST system where marked fish captured at the RST are transported upstream and released, or are new fish (e.g., hatchery) released upstream? If hatchery fish are used, have they already started to migrate from the hatchery or are they taken directly from the hatchery? Are the hatchery fish larger than the wild fish? If wild fish are used, are they held for a while before/after being marked/tagged. Are the fish released all on a single day at the start of the week, spread over multiple days, only during the day, or during all hours, etc.? Is the marking process traumatic enough to change the fish behavior while migrating down past the RST? These details are important because different release protocols and timing of movement from the release site to the RST change how missing data at the RST are imputed.

Has the RST protocol changed over time at each site, e.g., changes in hours of operation, crew, placement on the river, or the number of RSTs? There is a single mention in Chapter 5 on page 9, “If more than one RST is fished …”, so there may be variable numbers of RST used, and this will complicate fitting of a hierarchical model, which assumes homogeneity in protocol over time and sites.

Missing RST data for an entire week (i.e., the RST did not run at all during the week) breaks into two cases. If the missing data is completely random, i.e., unrelated to flow during fairly consistent conditions, the hierarchical model for capture-probabilities will “generate” plausible values for that week for the efficiency data. The spline model will generate plausible values for the run in that week (U_t) based on interpolating between the week before and that week. Uncertainties in the estimates of final run size will be larger compared to the case where the week is not missing.

Unfortunately, it is often the case that RST data is not missing at random, i.e., the RST cannot run at high flow levels, which is often associated with peaks of outmigration and low capture-efficiency. Plausible values for the capture probability can be generated from the efficiency model, assuming that flow was a covariate in the model and that the extreme flows are not too far outside the observed flow range. Generating a plausible model value for the run size in that week (U_t) is more problematic for the spline model. It is akin to guessing the height of a mountain based on the slopes as one approaches the summit. The BT-SPAS model allows for a "jump" in the run at known breakpoints specified by the user that allows the run in the week to suddenly jump at a high-low event, which is also useful when the data are complete. This feature has not been incorporated into BT-SPAS-X.

Partial missing data at the RST level is more complicated. Again, if the missingness is completely at random, i.e., unrelated to the size of the run or flow, then a simple expansion estimator may be appropriate. For example, if a RST only operated for 6 out of 7 days in a week when river conditions and the migrating run is fairly constant, then the number of unmarked (u_t) fish and recaptured fish (r_t) observed in that week could be expanded by 7/6. Depending on how quickly released marked fish migrate between the release site and the RST, some adjustments may be needed for the efficiency trial. Otherwise, the actual efficiency of the RST in that week is no longer properly described by the hierarchical model, and the expansion of the observed u_t will be too small. No information is presented in the report if this adjustment for trap efficiency has been made, except for a mention in the text of the next chapter (Chapter 5 Equation 8). If the missingness is not at random, there is no simple way to correct the data. The text under Chapter 5 Equation 8, indicates that, rather than adjusting to a 24-hour 7-day-a week schedule for the RST, values are adjusted to the mean number of hours the RST work. This may also work, but it again depends on how the missing times of the RST interact with the efficiency trials, etc.

Could the patchiness of the efficiency data also be driven in part by environmental factors? To what extent are environmental drivers also impacting the frequency and availability of efficiency trial estimates? For example, how many weeks are there in which the RST data are missing or only partially available due to high-flow events? This is a complicated issue and needs additional documentation in the report.

Bayesian Hierarchical model for RST efficiency data

The original BT-SPAS model assumed that sufficient efficiency trials were conducted each year so that only the year-site specific data is used in the spline model. Unlike the scenarios considered by BT-SPAS, efficiency trials at the mainstem RST are conducted sporadically, typically for only a few weeks each year. Efficiency data at the tributary sites is more frequent, but there are some sites-years when only a few trials are available.

The chapters require more preliminary data exploration of the efficiency data, such as summary tables of the number of trials by site-year, mean efficiency in each site-year using the pooled data to examine for the presence of year-specific effects; and variation in the raw efficiency around the pooled values to look for anomalous weeks.

The Bayesian hierarchical model (BHM) is described in Chapter 4 Equation 11a/11b and Chapter 5 Equations 1a/1b. The BHM is fit to the entire set of efficiency data (i.e., over all years and sites). A site-specific effect is fit allowing for sites to generally have consistently higher or lower efficiency across all years. A standardized flow covariate is fit, with flow standardized by the mean and standard deviation at each site. Finally, a random effect is added to model variation in the capture efficiency across years and weeks after accounting for site and flow effects. This final capture probability is then used in a binomial model for the efficiency trial.

The BHM is a sensible approach for the data, particularly in the mainstem, where recaptures are very small in virtually all trials, and for tributary sites with limited data. The BHM will shrink capture probabilities towards the mean efficiency in trials with very sparse data, as noted in Chapter 5, page 4. And, as noted later, the BHM provides a way to impute capture efficiencies for weeks in which no efficiency trials were conducted.

The BT-SPAS-X BHM does not incorporate a year-specific effect or a site-year specific effect (i.e., a random effect of year to account for consistent differences in capture-efficiency across weeks in a year across all sites or within a site). For example, the bottom panel of Chapter 5 Figure 8 appears to show that the observed efficiency trials tend to be lower than the mean (in red) for that site. A year-specific effect could be due, for example, to all sites experiencing a drier-than-normal year, with common impacts on efficiency from reduced flows. Year-site specific effects could be similar where a factor that impacts all weeks in a specific site-year occurred (e.g., forest fire).

The modeling of a common standardized flow effect makes a strong assumption that a difference of 1 standard deviation in the flow in a week relative to the mean flow over all years has the same effect on capture efficiency at all sites in one year. Sites may respond differently to changes in flow; this could be modelled by an interaction term between site and the standardized flow. Chapter 4 Figure 5 shows the modelled slopes across sites which shows that for most sites, the efficiency data likely does not have much contrast in the flow. This may be an artifact where efficiency trials cannot be conducted in very low or very high flows. When a model allowing for site-specific effects of flow was considered, Chapter 4 Figure 5, some impact was found for two sites, Butte Creek and Steep Riffle. Note the impact of what appears to be an outlier efficiency trial in the data for Butte Creek. However, the change for Steep Riffle may be an artifact of the very small contrast in flow at this site. These anomalous sites need further exploration.

Lastly, the fitted model assumes that unexplained variation in the capture probability has a common variation across sites and years. This requires further justification and exploration, e.g., Chapter 4 Table 3 shows how the mean efficiency varies across sites, the rationale for a site effect, but does not present information on the across-weekly variation.

The effects of fish size and fish type, wild or hatchery, were also explored. Data issues were encountered where not all efficiency trials recorded these covariates.

Chapter 4 Figure 6 shows low contrast in the mean fish size at release and so no effect of fish size was detected. However, this figure pools across years and a common year effect within a site may hide the relationship. There appears to be at least an across-site frontier to the data plotted here. Multivariate ordination may reveal a site by size interaction of some sort. This may imply that a model that uses mean fish size at release in addition to site-year specific effects may detect an effect of fish size.

Data issues were even more severe for the fish type, hatchery versus wild. The Upper Battle Creek trap site is adjacent to Coleman National Fish Hatchery, so hatchery-sourced fish are convenient there, especially when natural origin abundance and capture are low. Clear Creek rarely, if ever, relies on hatchery origin fish. The report needs to explain the source of fish in more detail. The MT found no differences in catch by fish type. Despite the data issues, the MT concluded in Chapter 4, page 27 “Capture probabilities for hatchery-origin releases in efficiency trials were approximately 50% lower than those based on releases of natural-origin fish (Figure 7).” Even if the modeling failed to detect an impact of fish type, this is very concerning because hatchery fish are frequently relied on when capture of natural origin is too small to provide the numbers for a decent efficiency trial. Lack of recording of type of fish in efficiency trials may be a data artifact (e.g., hatchery fish never used so never recorded), so some additional data review may be needed. Once the data review is concluded, a more complex model that includes fish type may be needed. Does the MT have any speculation on what would explain this observation? The trap provides a source of natural-origin fish by capturing actively migrating fish that are then tagged and returned upstream of the trap. Could the hatchery fish be exhibiting a different behavior? Not migrating? Higher mortality?

These more complicated models need to be examined (e.g., adding more covariates to produce more reliable predictions is suggested in Chapter 4 Section 4.1.2) and a rationale provided on why the simpler model is used. The effect of not fitting a year-specific or year-site specific effect is noted on Chapter 5 page 5 and could lead to “biases” in estimates of efficiency for certain years. For example, the bottom panel of Chapter 5 Figure 8 may indicate that for this site-year, capture-efficiencies tend to be lower than the grand mean. Similarly, because of hydrological and site-specific differences, efficiency may vary with flow differently each year.

The investigators highlight the low capture probability of marked juvenile Chinook salmon at the two mainstem RST. The mean of observed capture probabilities at the Knights Landing (n = 139) and Tisdale (n = 47) RST sites were 0.47% and 0.25%, respectively. These capture probabilities are approximately 15-fold lower than the estimated mean for Sacramento River tributaries of approximately 2.5%. Indeed, Chapter 5 Tables 2a/2b show the presence of many efficiency trials in the mainstem, where a small number, or none, of fish were recaptured. This may cause difficulties for the BHM. First, the model equation attempts to fit each trial's capture efficiency on the logit() scale. Unfortunately, the logit of 0 is -∞ which causes difficulty for most computer code. The underlying code may "converge" to using a value, say of -100, which could cause serious negative bias in the overall mean leading to a significant negative bias in the imputed capture efficiency and significant positive bias in the estimated weekly abundances. This also could give rise to the very large uncertainties in the capture efficiency estimates as seen in Chapter 5 Figure 5 (lower panels). Seber (1982) shows that significant bias can be introduced when the number of recaptures in a capture-recapture experiment is below 6. This is because $E[\frac{1}{\hat{p}}] \neq \frac{1}{E[\hat{p}]}$. Both issues may explain why the mainstem estimates of abundance are so high. It may be beneficial to try the standard "trick" of adding 1 to the counts of releases and recaptures to avoid 0 estimates and reduce bias from a small number of recaptures.

The MT should also encourage field crews to use more RSTs at the two mainstem sites so they sample a volume of water and proportion of the salmon migration that is consistent with the tributary sites and at Red Bluff Diversion Dam where up to five RSTs are deployed.

Bayesian model for weekly and overall abundance

BT-SPAS-X uses a similar approach as BT-SPAS to fit a spline curve with two distinct changes:

1. First, unlike BT-SPAS, the efficiency data may be sparse for a particular site-year. Consequently, imputed values from the hierarchical model for the capture-probability over all sites and years are used rather than just efficiency data for that site years (Chapter 5 Equation 9 and Chapter 5 Figure 2).
2. Second, like BT-SPAS, a spline with constrained random variation about the spline is used to model the run shape. The spline ensures that sparse data do not lead to large changes in abundance from week to week that are strictly data artifacts. Additionally, the spline is used to impute the weekly abundance when the RST was not running. But unlike BT-SPAS, a flow covariate is added to the spline function. BT-

> SPAS allows for “jumps” in the outmigrating weekly run size and does not use flow as a covariate.

Conceptually, the number of unmarked fish captured by an RST in week (u_t) is expanded by the imputed capture efficiency ($\hat{p}_t$) to estimate the weekly abundance as $\widehat{U}_t = \frac{u_t}{\hat{p}_t}$ but smoothed by the underlying spline.

Because the efficiency data is “imported”, it does not have to be estimated based on the number of unmarked fish captured each week. This is the reason for the “cut()” function mentioned in Chapter 5 Equation 9 and Chapter 5 Figure 2.

The dependence of weekly abundance on flow (BT-SPAS-X) or allowing discontinuities in the spline (BT-SPAS) both try to deal with cases where the weekly abundance suddenly jumps, for example, after a rainfall event which then increases the flow and “induces” migration. High flows also make it likely that the RST must be stopped due to safety reasons, so the peak of the run often occurs when there is no or limited data! Both approaches try to deal with the difficult problem of estimating the peak of the mountain based on how fast the mountain is rising without actually climbing the mountain.

The use of the flow covariate in BT-SPAS-X for estimating weekly abundance assumes that increased flows have the same impact across all parts of the run (a linear response to flow on the log scale). But in flashy systems, there could be several instances of increased flow. Fish may respond more to the first flash event than to latter events (for example, the majority of fish could move on the first flash event). BT-SPAS allows unspecified sizes of jumps at flow events, but not all flow events need to have the same impact.

However, this is moot because the MT states (Chapter 04 page 22) “The current version of BT-SPAS-X does not include the $\phi X_{s,y,w}$ [the flow covariate] element in Equation 13.” i.e., the current model does not include using flow to have an impact on abundance. This implies that BT-SPAS-X behaves like BT-SPAS without the ability to accommodate sudden jumps in abundance. One consequence is that the estimated variances of error terms around the spline model are larger than they need to be, which implies that estimates of weekly abundance are more variable than they need be (see discussion at end of Chapter 4, page 22 and top of Chapter 4, page 23). This increased variability around the spline also implies that the BHM may estimate a very large weekly abundance which could inflate the overall abundance. This is the rationale for capping the weekly run size at a large value (see Chapter 4 Equation 4).

Adjusting for the proportion of spring-run juvenile Chinook salmon

Weekly estimates of spring-run Chinook salmon abundance were derived from total Chinook salmon abundance using PLAD model proportions, which captured strong seasonal variation and generally provided precise weekly estimates, particularly when spring-run Chinook salmon dominated the catch. After adjusting for the proportion of spring-run Chinook salmon in each week, the total abundance of spring-run Chinook salmon is found by adding the weekly adjusted abundances. The Bayesian model automatically propagates uncertainty from each week to the final grand total.

With respect to the mainstem sites, the average CV in annual abundance estimates (all run types combined) across years was approximately 30% (Knights Landing: 18–51%; Tisdale: 19–46%). Generally, annual abundance estimates at mainstem RST sites were highly uncertain.

Please refer to our comments on MT Chapter 6 about the use of the PLAD model to identify the proportion of juvenile salmon by run type.

Submodel Results

Annual abundance estimates had an average coefficient of variation of 26% (tributaries) and 37% (mainstem) while precision varied across sites and years depending on efficiency trial coverage and weekly catch patterns (e.g., Chapter 4 Figure 8; Chapter 5 Tables 4a/4b).

This mainstem model’s estimate of SRJP has much higher CVs than those from the tributary model. Consequently, the investigators concluded that the utility of Sacramento River mainstem abundance estimates for calculating a spring-run JPE may be low relative to tributary-based approaches owing to the large uncertainty in the estimates. Also, they suggested it was unlikely that the uncertainty of mainstem abundance estimates could be reduced by conducting more-frequent efficiency trials unless capture probabilities or the ability to estimate capture probabilities can be improved (see above).

Annual abundance estimates were highly correlated among the two RST sites across the overlapping set of years (n = 12, $r^2 = 0.97$). It would be useful to also examine correlations in weekly abundances between the two sites. This could identify weeks where the fit may be problematic (see above). Given this high correlation, the MT should also consider how the two RST sites could be modelled together, or how one can be used to estimate the other if sampling weeks are missed.

Estimated total Chinook salmon and spring-run Chinook salmon juvenile abundances at the mainstem traps and combined RSTs should be directly compared, as it appears that the mainstem RSTs provided much greater abundance estimates than the tributary RSTs even after high mortality in the mainstem (e.g., based on Chapter 5 Table 4 comparison with Chapter 3 abundance values). The MT noted this discrepancy during an introductory meeting with the Review Panel. The discrepancy is also highlighted in the migration timing chart in Chapter 3 Figure 18. Part of the discrepancy may be due to the issues in modeling very low capture-efficiencies with many 0 recaptures.

Investigators noted that alternative approaches to estimate capture probability at mainstem RST sites, such as paired coded-wire tag (CWT)/acoustic tag releases, could lead to reductions in the uncertainty of abundance estimates. The problem with acoustic tags is that they can only be inserted into relatively large juveniles (approximately 80 millimeters or larger). Much smaller “Shad Acoustic Tags” or PIT tags could potentially allow estimates of survival and migration rate for much smaller salmon, e.g., fish as small as 45 to 50 millimeters (See more information in Chapter 9 and in Charge Questions).

Overall, the statistical models in the chapters are most appropriate and appear to be implemented properly. These chapters could be improved by:

1. Showing plots of the fitted spline curve (on log-scale) versus the estimated weekly abundances similar to Chapter 5 Figure 5. This will allow the reader to assess if discontinuities in the spline are warranted.
2. Additional model assessment as outlined in section 0 of this report.

3. Comparison of results from BT-SPAS-X to simple pooled-Petersen estimates to identify any anomalous site-years.

Recommendations

Combine chapters 4 and 5 into a single chapter, as the modeling framework is common to both.

Provide additional information on sampling protocols

Additional information on the sampling protocols at the RST traps are needed. For example, how is data adjusted for RST not running 24-hour 7-day-a week. At what flow levels are RST pulled for safety? Is the same number of RST used in all weeks?

How are efficiency trails conducted? For example, are all the fish for a trial released in one day or are the releases spread over the week. Do most fish travel between release and the RST in 1 day or is movement spread out?

Examine RST data protocols to verify that fish type being released is actually unknown and not just “not recorded” to improve the modeling of capture probability as a function of wild/hatchery origin.

Provide some raw data summaries/analyses

Preliminary data summaries of RST data are needed, such as what fraction of the week are the RST running? When are RSTs not running in relation to flow. Plot the raw efficiency data by various covariates such as flow or time or year to provide insight in the form of the hierarchical model.

Additional modeling required.

Plots of the efficiency data show that year-specific and site-year-specific effects may be warranted in the hierarchical model. BT-SPAS-X should allow for discontinuities (jumps) in the spline to accommodate sudden increases in the run due, for example, to high rainfall events. There are about 150 site-years of data that would have to be examined, so for an initial approach, an algorithmic approach may be helpful to flag potential site-years for further examination. For example, a jump may be warranted when the number of fish captured at an RST in week t+1 is, say 10 times larger than in week t. Or sudden increases in flow might also be an indicator of a need for a discontinuity. Or both indicators could be used. Some human oversight will also be needed.

Explore mainstem RST models where data is adjusted, by adding 1 to the number released and recaptured, to avoid many 0 recaptures and the biases introduced by a small number of recaptures.

Adjustments to protocols in future years

How much flexibility is there in the scheduling of efficiency trials? How is it decided when to conduct efficiency trials? Presumably, trials that occur when there are many outmigrations lead to the most precise estimates. If flow is to be used as a covariate, efficiency trials should be attempted across a range of flow values to improve contrast.

Chapter 6 PLAD Run-assignment Model

Chapter Overview

The probabilistic length-at-date (PLAD) model was developed to address the major challenge of identifying the run type (spring, summer, fall) of juvenile fish detected at RSTs. This PLAD model is general enough to use both past and future data, and to provide both the probability density function of the likely run-type distribution at each site and the probability of a specific run type for a known-length fish captured on a known date.

While length at date (LAD) models have been used in the California Central Valley for several decades (Harvey and Stroble 2013; Brandes et al. 2021), they assign a categorical run type deterministically, using static size and timing boundaries to classify juveniles. Although widely applied, LAD models are known to have high misassignment rates. Genetic run identification methods have improved substantially and are increasingly used, including targeted genetic sampling implemented for the SRJPE Program; however, genetics cannot be applied to all sampled fish and is unavailable for most historical datasets. To address these limitations, PLAD uses genetically identified fish to estimate probabilistic predictions of run type and associated uncertainty for non-genetically tested juveniles, using covariates such as fish length, capture date, and location. PLAD computes a set of probability density functions for the individual's fall in each run type, thereby moving the LAD model to a PLAD (probabilistic length at date) model. This moves from a LAD framework to a PLAD framework allows the MT to incorporate uncertainty in run type into management models, which is important since the SRJPE relies on accurate assignment of spring-run fish. The MT also addressed the possibility of uncertainty arising from specific scenarios and the need for users to change their choices in model customizations based on the application, which grounds models in realistic ecological scenarios. PLAD models were tailored to individual RST sites in the Sacramento River and its tributaries and then applied to historical monitoring data. When coupled with the BT-SPAS-X model (Chapters 4 and 5), PLAD-run assignments enable hindcasting and forecasting of SRJPEs.

The MT developed PLAD to meet two main goals:

1. to assign run type probabilistically instead of deterministically, and
2. to create a framework that would allow for spatial variation.

First, the MT developed a finite mixture model of run types based on fork length, date of capture, and location of capture. Location was used to create site-specific models, while the mean and variance of fork lengths on a specific date were used to determine the probability that a fish was a fall-, winter-, or spring-run type. Any specific day will therefore have a different mix of fork length distributions based on the combination of runs going through that RST site on that date (See Chapter 6 Figure 1). Modelers developed two site-specific estimates:

1. the change in mean fork length over time using a log-linear growth model and
2. the change in run proportion over time.

Including time as a covariate in both models reduces the overall number of parameters. Parameters in the PLAD model are estimated using Bayesian methods based on field collected data from two main data sets: the genetic identification of run type for a sub-sample of fish from each RST site, and the shape of the catch distribution at that site. Fork lengths from genetically identified fish with known run type were used with a lognormal likelihood to get information on the mean and variance of the distribution. Non-informative priors were used for all parameters except the daily growth rate, which was assigned a prior based on estimated juvenile Chinook salmon growth from previous LAD work in the California Central Valley region (Harvey et al. 2014). The MT presents an example of the model-fitting process for a single tributary that has all three run types, Battle Creek, showing that winter-run fish emerge earliest for a given brood year and grow slowest and fall-run emerge latest and grow fastest, with spring-run between in both emergence timing and growth rate. Given the low abundance of winter-run Chinook salmon, the spring-run typically has dominant proportions earliest in the migration season, emerging early and tapering off, while the fall-run is dominant later in the migration season, emerging later and increasing in proportion as the season progresses.

In their current state, the PLAD models are based on three years of data, 2022 to 2024, highlighting the issue of how year-to-year shifts in relative abundance and growth rates of each race affect PLAD findings when there are no genetic data (i.e., prior to 2022). Run-specific size at a given date is key to the prediction but run-specific fish length will vary from year to year depending on timing of emergence of fry, food availability and temperature. The MT tested this issue with Battle Creek data from 2025 and found that the model tended to accurately predict spring-run Chinook salmon, while winter-run fish were misclassified as fall-run fish. The MT noted that the misclassification was likely due to a change in stock composition between 2025 and 2022 to 2024. They also noted the uncertainty of PLAD predictions increased late in the migration season. In contrast, juveniles captured early in the migration season had low uncertainty in their run assignment. Importantly, the MT recognized the issue of year-to-year growth variability and has recommended development of PLAD models in a hierarchical fashion while noting the need for additional years of genetic data.

Key Limitations

The Review Panel identified the following key limitations in the PLAD modeling framework:

1. PLAD has relatively high data and parameter demands, as it requires sufficient temporal coverage and sample sizes to reliably estimate time-varying length distributions.
2. Parameter identifiability can be weak when cohorts overlap in time, and

extrapolation beyond the observed date range may be unreliable.

3. Model accuracy depends heavily on the amount and quality of genetic run identification data used to calibrate the model, such that distinguishing runs becomes harder downstream where size differences narrow (Brandes et al. 2021).
4. Model performance can be sensitive to assumptions about growth variance and distributional form (e.g., normality), which may not hold if growth is strongly size-dependent, episodic, or influenced by unmodeled environmental drivers.
5. Potential to confound growth processes with selective sampling or migration timing, making it difficult to disentangle whether changes in observed length-at-date reflect true growth, size-selective entry, or gear selectivity.
6. Assumptions about growth patterns and distinct run timing may not hold in all years or locations, contributing to uncertainty in run identification (Nelson et al. 2022).

Recommendations

PLAD model uncertainty and robustness

The difficulty of this mixture problem can lead to substantial uncertainty in estimating the proportion of spring-run Chinook salmon. There are two types of robustness to examine: robustness to data availability and robustness to the assumptions made during the modeling process.

The Review Panel recommends exploring additional data types that could improve estimates and increase robustness. One reason for high overlap in uncertainty between spring and fall-run types is limited sample availability on certain dates, leaving run assignments uncertain early in the season for Butte Creek, Lower Clear Creek, Mill Creek, and Deer Creek. A major driver of uncertainty in this model is a lack of samples, so the Review Panel recommends exploring other data sources or alternative collections to supplement available data.

In addition to increasing the overall availability of samples, another way to address the issue is to optimize sampling times. Chapter 6 Figure 4 shows when fish were sampled. To fit run curves like the hypotheticals displayed in Chapter 6 Figure 3, teams should sample near the peak of each run and in the tails. While the spread of sampling dates appears to be adequate, some investigation is needed on optimizing the mean date of sampling, i.e., can there be a shift of possible sampling from later in the migration year to earlier in the migration year for each run type? This can be investigated using simulation studies. For example, from Chapter 6 Figure 3, the early-run curve should be sampled near the peak and near the tail; the late run curve should be sampled near the left side and at the peak, etc.

In addition to robustness to field collection availability, the panel recommends exploring PLAD's robustness to model assumptions. The model makes strong assumptions of homogeneity of catchability across genetic samples and species, and across different lengths. The MT should consider the following: how robust is the PLAD model to assumptions of homogeneity of capture across time/species and to growth patterns across time? How robust is PLAD to the unimodal shape of run curves? The same run timing curve is assumed to occur over all years, indicating a lack of interannual variability; is this an appropriate assumption? The panel recommends exploring additional covariates, for example, the peak of the run curve may depend on when the first high flow event occurs, or a random effect could be added that allows the run curve to shift left or right by year depending on environmental conditions. Simulation studies should be conducted to investigate the impact of assumption violations on the resulting estimates.

Continue genetic testing to meet PLAD data needs

Genetic data will be needed annually to refine PLAD estimates at RST sites, especially those that rely heavily on data from years prior to genetic sampling. In addition, the increase in winter-run fish during 2025 indicates that continued genetic testing and potential future updating of the PLAD model is necessary to account for changes in run type as future conservation efforts and management actions go into effect.

Test other model parameterizations

The PLAD model is also vulnerable to assumptions around parameterization types. The panel recommends that the modeling explore a few alternative versions of the PLAD model to assess the sensitivity of run proportion predictions to assumptions about the shape and type of the distribution and the growth model. For example, the MT could test a version in which the β_i

parameters are the same across different runs. The PLAD model assumes a unimodal curve for each stock. How much bias is introduced if there are two peaks of migration? This can be investigated through simulation and shown using changes in the ρ terms.

In addition to further simulation testing, there are a few areas that simply need more discussion and clarification. For example, the RST catch data shows total composition of the run and therefore isn’t resolved into run type. How are the posterior samples on run size from RST used?

Site specific concerns and increased data collection

High uncertainty in early migration season estimates is a common issue across creeks, as is an overlap in timing between spring- and fall-runs. As discussed above, increased data availability improves models, for example Upper and Lower Feather. Is increased data collection an option for sites with fewer data like Deer Creek?

One result of the high uncertainty is that fall-run fish appear very early in some sites (i.e., Yuba and Butte creeks); is this realistic? If not, can the MT suggest whether this uncertainty biases the final SRJPE estimates, and if so, how?

Is it reasonable to assume that juvenile growth parameterization will remain the same across time? How often will the PLAD model be updated? Perhaps a “moving window” should be used once many years of data are available, i.e. only use data collected in the 5 years up to and including the year of interest to develop a PLAD model.

The MT also suggests that the use of hierarchical modeling can be used to group RST sites according to shared spatial domains. The Review Panel believes that this will be a difficult task, because it is hard to determine which sites are “spatially close”. Presumably, the sites are in different watersheds. More information is needed on the grouping determination (for example, will they be grouped based on geomorphology?).

Finally, Feather River presents an issue in this chapter as it does in others, as it is one of the highest contributors to overall production. How will uncertainty in PLAD propagate forward into the overall SRJPE?

Chapter 7 Stock-Recruit Model

Chapter Overview

This chapter is a standard analysis of stock-recruitment (S-R) data using the Ricker model. The analysis links spawner counts of spring-run Chinook salmon in Sacramento River tributaries to annual juvenile spring-run outmigrant abundances in those tributaries and in the mainstem. S-R analyses are needed to initiate juvenile forecast models in winter by predicting average outmigrant abundance from a given spawner level.

Annual juvenile abundance estimates are produced from BT-SPAS-X (Chapters 4 and 5) and PLAD models (Chapter 6), while spawner abundance is estimated using a variety of survey types such as passage counts, redd counts, adult swim surveys, and carcass surveys. Details of these surveys were not described in the MT report. But in some tributaries, inconsistencies in annual survey methods and lack of sampling in some years has reduced sample size, years of data, which can limit the number of covariates (e.g., flow, temperature) in the S-R model and reduce contrast needed to show the relationship between spawners and their progeny. Environmental factors like flow and temperature are tested to account for potential effects on survival rates during upstream migration, holding in the river, spawning and incubation, and early rearing phases.

Multiple S-R models were evaluated, including null models, discrete water year type models, and continuous covariate models incorporating temperature and flow variables affecting survival during migration, spawning, incubation, and early rearing. Model performance was assessed using goodness-of-fit metrics, process error, covariate reliability, and leave-one-out cross-validation. Across sites, the best S-R models explained a moderate amount of variation in log outmigrants/spawner (r^2 approximately 0.4 to 0.7) and had relatively high out-of-sample error in predicting outmigrant abundance (relative errors 50 to 100%).

It is not uncommon for S-R models to exhibit high variability, and the Ricker S-R model incorporates observational error in the number of recruits. The Ricker S-R model allows testing for density dependence, which was observed here in a couple of cases (Yuba River, Knights Landing). The Ricker S-R model is also a common stock-recruit model choice for modelers planning to include environmental covariates as it can be linearized fairly easily. However, the limited number of data years constrains most of the models to only a single environmental covariate. Nevertheless, multiple models are developed using a variety of covariates, but only one covariate per model.

Temperature and flow covariates improved model fit in some watersheds, e.g., reduced productivity under warmer temperatures and, in several cases, higher productivity with higher flows and wetter water years. More review by the MT of the results in Chapter 7 Table 3 is warranted. For example, in most cases, models with *rr_mean_flow* (mean flow November to July), *rr_max_flow* (max flow November to July), *rr_med_flow* (median flow November to July), and *rr_min_flow* (minimum flow November to July) gave similar *MAE_rel* (out-of-sample error), except for Chapter 7 Tables 3b and 3f. Why? Is there an "outlier" among these covariates for one year that is driving the results? These results are surprising given the high correlation among these covariates shown in Chapter 7 Appendix A. Water year did not help explain recruitment in many of the creeks.

More discussion by the MT is needed on how they will use models having different covariates. For example, in each creek several covariate models have similar relative median absolute error values, so how would a user choose,

1. whether to use the *si (*spawning-incubation) models, the *rr* (rearing) models or the *mi* (migration) models, and
2. within a life stage, how should users choose which covariate model to use if several perform similarly well?

Is there a way to combine flow and temperature into a single variable (see comment below)? Many of the creeks showed the influence of flow during emergence to outmigration (*rr*) and of temperature during the spawning and incubation (*si*) phases. Future years of monitoring should help the MT include more covariates in a single model.

A common tactic in the face of many covariates is to conduct a Principal Component Analysis to investigate what form the first component takes. This may indicate that a composite variable combining several covariates is suitable, such as adding z-scores from several variables. However, we note the limitations of this approach previously.

If a likelihood framework were used, it would be relatively straightforward to use information-theoretic methods (such as AIC) to combine results from multiple models, where each model is given a weight that accounts for complexity and fit. This would also fit more complex models (e.g., multiple covariates). However, the increase in fit from more covariates may not overcome the fitting penalty from more covariates, and so, more complex models may not be suitable. Multiple model averaging is much more complicated for Bayesian models (e.g., reversible jump MCMC), but simple approximations are available via Bayesian Model Averaging (BMA; Hoeting et al. 1999).

Another concern is the partial treatment of the error-in-variables problem where there is uncertainty about both the recruitment, from the BT-SPAS-X models, and the spawning population from redd surveys, etc. The measurement error in recruitment is addressed by incorporating a component into the fitting process. However, measurement error in the spawning numbers has not been considered because no error data were available. Ludwig and Walters (1981) considered this problem and showed that errors in the measurement of spawning stock are particularly problematic and lead to a specific type of statistical bias. While measurement error in spawning numbers is difficult to quantify, the analysis should investigate different levels of measurement error (e.g., 20% or 50%) to see how the results change relative to assuming no error.

More details on the model assessment of the fit are needed. The current analysis uses R^2 and the process error, but the MT should also consider trace plots, Rhat, predictive posterior plots, autocorrelation of model residuals, etc.

Although the MT identified several issues that have led to uncertainty in spring-run outmigrant forecasts, they suggested that consistent, robust monitoring of spawners and outmigrants each year could improve S-R relationships and forecasts in the future. With additional years of data, additional covariates could be used in the same model.

Recommendations

The MT described a variety of data limitations that influenced the relationships. For example, in one tributary, spawner numbers were incomplete because one spawning reach was not consistently sampled each year, such that the S-R could not be performed. The number of years of available S-R data was a consistent limitation, especially when attempting to incorporate more than one environmental covariate. The S-R relationships were highly variable or non-existent, in response to error in estimated abundance of outmigrants (PLAD applied to RST counts), error in spawner counts, and/or unexplained density independent variability.

Evaluate and refine covariate use

The MT should continue to search for explanatory variables. Flow- and temperature-based covariates led to only modest improvements in model fit, and in many cases did not reduce out-of-sample error in outmigrant abundance predictions, the exceptions being Deer and Butte creeks. The MT should explore refined continuous temperature and flow covariates, while recognizing site-specific variability and potential masking effects from measurement error or limited environmental contrast.

For mainstem S-R relationships, peak flow may be less indicative of extended rearing season conditions than water year index, or thresholds of flow that breach various flood control weirs and engage high-value habitat in the flood bypass reaches.

Identify other discrete models & models that combine covariates

While water year was not ubiquitously successful as a covariate, is there another way to combine flow and temperature into a single model? Many of the creeks showed the influence of flow during the early rearing phase, emergence to outmigration, and of temperature during the spawning and incubation phases. For example, the time of emergence combines information on temperature and developmental effects; it could allow for an implicit interaction effect without needing to include both covariates.

Applying high-performance models or weighting tributary contributions

Because Butte Creek has relatively strong S-R model performance and is a major contributor to overall production, it may serve as a potential proxy for system-wide dynamics. Using Butte Creek as a reference could provide redundancy during data gaps or catastrophic events and may increase user confidence due to its stronger model fit. Correlations between tributary recruitments and spawners should be explored to determine whether common environmental variables are at play.

However, differences in habitat, hydrology, and population structure across tributaries should be carefully considered before generalizing results. Alternatively, recruitment estimates could incorporate a weighting scheme that places greater emphasis on tributaries with stronger model fit and more reliable data.

Model selection across life stages & covariates

When multiple models perform similarly within each creek, users must decide which life stage (spawning–incubation, rearing, or migration) and covariate model to apply. Perhaps selection should prioritize biological relevance, data availability, and management objectives, rather than relying solely on small differences in model error. If performance is comparable, models that represent key limiting processes or are supported by stronger ecological justification may provide more useful decision support. Alternatively, the MT might average the outputs from the most appropriate models, or conduct a Principal Component Analysis as described above.

Reduce and address uncertainty in spawner abundance data

An important issue identified by the MT is the lack of error estimates associated with spawner counts in each tributary. One potential source of error is the temporal and spatial overlap of spawning spring-run, fall-run, and winter-run Chinook salmon. PLAD was used to estimate run-specific outmigrants, but there was no mention of whether the runs of Chinook salmon overlap on the spawning grounds, potentially leading to error in spawner counts for spring-run Chinook salmon. Another source of error is the degree to which the female-to-male Chinook salmon ratio varies from year to year. For example, in some Alaskan watersheds, female Chinook salmon are much less abundant than male Chinook salmon because females mature at an older age to gain fecundity, and have lower survival. Also, redd counts are multiplied by 2.5 to account for more males than females on the spawning grounds. Analyses should consider this ratio when possible.

Tributary-specific differences in spawner enumeration methods can also introduce inconsistencies that can affect S-R model results. For example, Knights Landing relied only on upstream passage estimates from Battle and Clear creeks, and adult abundance from Butte, Mill, and Deer creeks may be underrepresented, potentially biasing productivity estimates. Multiple sections mention that Knight's Landing values are inflated due to the lack of information about spawner abundance for Mill, Deer, and Butte Creek. This is an important issue because the outmigrant/spawner estimate will be overestimated if spawner counts are used from only 2 of the 5 tributaries. Approximating missing spawner counts could be attempted, but this would increase uncertainty, which should be incorporated into the model output.

Methods to improve the accuracy of spring-run Chinook salmon spawner counts involve:

1. Consistently monitor each spawning reach every year
2. Stock identification (fall v. spring, hatchery v. natural)
3. Survey life of redds, spawner counts, & observer efficiency
4. Document female vs male spawners; do not assume 50:50

Chapter 8 In season-Outmigrant Model

Chapter Overview

This chapter describes the in-season outmigrant model that forecasts the timing and abundance of spring-run Chinook salmon juveniles passing RSTs. The in-season outmigrant model answers two questions:

1. when do juveniles go through the RST sites, and
2. how many juveniles go through each RST site?

Using multi-year weekly outmigration data, the model predicts the proportion of fish passing each week during the migration season (i.e., November 1 in year *t-1* to October 31 in year *t*) and applies current observations to estimate remaining seasonal and total annual abundance when full-year data are not yet available. The model estimates the outmigration by proportions (e.g., the fraction of runs that pass each week) and by percentile (e.g., the date when 50% of the runs migrated). Juvenile abundances are estimated as a proportion of the total number of outmigrants, so that if sufficient data are present for previous weeks in an outmigration season, the abundance of juveniles can be predicted for each subsequent week in the season. Given an observation of outmigrant abundance up to a specific week in a forecast year, the outmigration-timing model can predict the total outmigrant abundance for the year. Weekly RST predictions are routed to the Sacramento–San Joaquin River Delta using a separate model that accounts for travel time and mortality. These routed estimates are combined across tributaries to produce a spring-run juvenile production estimate in the Delta.

The primary relationship in the in-season outmigrant model is between that of the median run time (*phi*: the mean of the beta distribution used in this model) and the kurtosis of the overall run time distribution (*lambda*: the spread/peaked-ness of the distribution around *phi*). The use of a beta distribution avoids “fixed” points, allowing greater flexibility at the cost of increased complexity (this report Figure 3). The in-season outmigrant model was applied to spring-run juvenile abundance estimates from RST sites, six of which are in tributaries and two of which are on the mainstem of the Sacramento River. The model generally fit cumulative weekly outmigration data well and captured site-specific differences in timing. Two states emerge between the creeks, with earlier median run time creeks also exhibiting narrower run time distributions and later run time creeks exhibiting flatter run time distributions. Mixed effect models with a tested flow covariate are used to estimate these two parameters (*phi* and *lambda*). Upper Clear Creek exhibited the earliest and most rapid outmigration, while Mill and Deer creeks showed later and more gradual patterns. A negative relationship between median outmigration timing and run-timing steepness was observed at most sites, although correlations were generally weak and uncertain except at Battle Creek.

The correlation between *phi* and *lambda* is represented by a joint hyper-distribution that takes a multivariate normal form based on the estimated across-year means of the two variables. If no correlation is present, the Pearson’s coefficient between these two variables will be near 0 and the model will not force a correlation to be used. The model is first fit to historical data at an RST site, then used for forecasting in a Bayesian framework. Correlated random draws of the intercepts for *phi* and *lambda* are obtained from the posterior of the hyper-distribution parameters, allowing interannual variability to be represented. The posterior for beta is used to modify the covariate value for that year and thereby adjust *phi* or *lambda* when fixed effects are present. Deviates are then generated for each week in the forecast time period using a posterior for the process error. The resulting total abundance in a year is the log-transformed cumulative weekly abundance for a specific week from BT-SPAS-X, with simulated error drawn from a normal distribution. This total abundance is then used in the forecasted weekly estimate, where it is multiplied by the proportion of the run that passes in that week to obtain an abundance value for that week at that site. As the season progresses, more fish have passed through each RST site, reducing uncertainty as data availability increases.

Forecast uncertainty was largely driven by interannual variability and was highest early in the outmigration season when only a small proportion of the population had passed RST sites. There was considerable variability in the forecasted total outmigrant abundance, especially when the forecast was made early in the migration season (late December and early February). This was expected because a relatively small proportion of the migration had passed the RST at these times. Forecast precision improved later in the season as more fish were observed, resulting in lower in-sample and out-of-sample forecast error across most sites. Most forecasts improved by early March, with relatively little improvement seen between later dates (March 1 to 29). The exception to this is at the sites with early-migrating populations, which have more accurate forecasts earlier in the season than late-migration populations because a large proportion of the run has already passed the RST when the early-season estimates are made. Upper Clear Creek, for example, had consistently low forecast error due to early migration timing.

Several other questions are posed and answered by the MT as it chooses how to incorporate various data sources and terms into the model. Two that the MT demonstrate in both mathematical methods and with model validation results are related to correlation:

1. can flow be used as a covariate to improve run time and abundance estimates through the correlation of flow and outmigration, and

2. can a lag-1 term be used to improve modeling of extra-beta variation (i.e., can variation estimates be improved by allowing weekly deviates to be correlated with previous weeks)?

The MT concludes that the answer to both of these questions is no; that

1. correlation values between flow and *phi* are both low and uncertain, a combination that makes this particular covariate untenable as a predictor at most of the RST sites, and that
2. correlation coefficients are so low (close to 0) that the lag model is not even presented as part of the chapter and is instead only shown in the appendix (thus the weekly deviates can be randomly pulled from the normal distribution in Chapter 8 Equation 2a instead of calculated with autocorrelation to previous weeks).

Models incorporating peak flow covariates produced results similar to null models, with only minor differences during early forecast periods. Peak flow covariates had little influence on timing, with only minor and uncertain effects observed at the Yuba River site.

Recommendations

Proportion Analysis: A proposed alternate approach

An alternative to the current method of fitting a beta distribution is to select a subset of weeks for proportion analysis and a subset of percentiles for timing analysis, with linear interpolation as needed. This reduces the complexity of the previous approach to a more tractable problem.

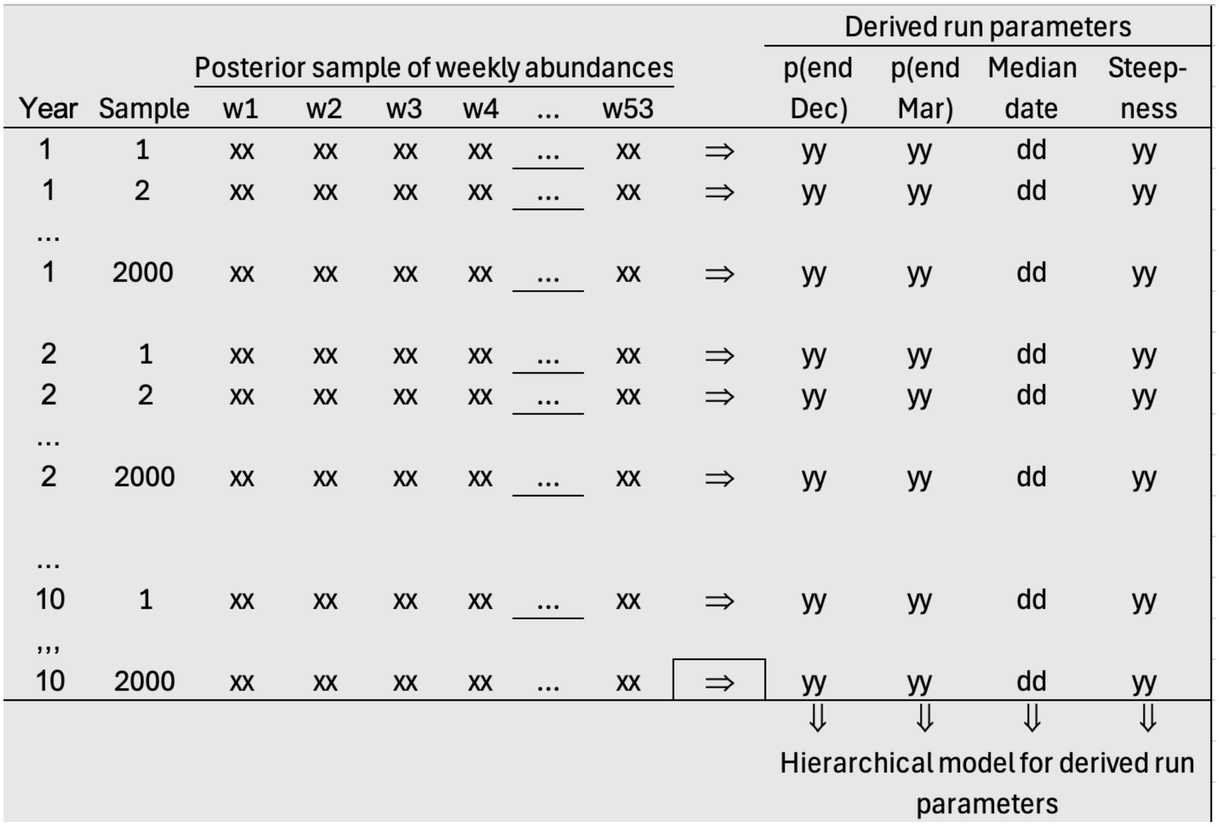
The MT originally attacked this problem by fitting a beta distribution to the cumulative proportion of the run and then using it to forecast various parameters of interest. This method is overly complex and does not integrate the posterior data effectively. The MT mentions that the fitting methods are difficult to explain and update, and lack flexibility.

Characteristics of run timing (e.g., proportion of the run that passes by a specific week) are derived from posterior data on the abundance of the run that passes the RST in each week. In much the same way that total abundance is a derived variable of the weekly run abundances (sum of the abundances), proportions by week and percentiles are also derived variables and will be treated similarly. For example, it is relatively simple to modify BT-SPAS-X to compute these for each year that is analyzed at a site (BT-SPAS does this).

The key issue is how to handle the 10 years of data simultaneously. Instead of building a super-model to handle all 10 years, the MT could use the posterior values from each of the 10 fits as a “pooled” posterior sample (this report Figure 3).

Worked example of the proportion analysis:

Suppose there are 10 years of data from a site and that BT-SPAS-X is run for each site. The final product of BT-SPAS-X is a sample from the posterior distributions of the number of fish that pass the RST each week (say, a sample of size 2,000) for each of the 10 years the site was in operation (this report Figure 3). This posterior distribution must then be manipulated to yield the posterior distribution of the run timing over the 10 years of data without analyzing all 10 years simultaneously.



		Posterior sample of weekly abundances							Derived run parameters			
Year	Sample	w1	w2	w3	w4	...	w53		p(end Dec)	p(end Mar)	Median date	Steep-ness
1	1	xx	xx	xx	xx	...	xx	⇒	yy	yy	dd	yy
1	2	xx	xx	xx	xx	...	xx	⇒	yy	yy	dd	yy
...												
1	2000	xx	xx	xx	xx	...	xx	⇒	yy	yy	dd	yy
2	1	xx	xx	xx	xx	...	xx	⇒	yy	yy	dd	yy
2	2	xx	xx	xx	xx	...	xx	⇒	yy	yy	dd	yy
...												
2	2000	xx	xx	xx	xx	...	xx	⇒	yy	yy	dd	yy
...												
10	1	xx	xx	xx	xx	...	xx	⇒	yy	yy	dd	yy
,,,												
10	2000	xx	xx	xx	xx	...	xx	⇒	yy	yy	dd	yy
									⇓	⇓	⇓	⇓

Hierarchical model for derived run parameters

Figure 3. Sketch of estimation run timing values from multiple years of data.

Consider, for example, estimating the proportion of fish that pass the RST by the end of December. This is done by fitting a simple hierarchical model as follows (this report Figure 3). It is easy to add covariates to the hierarchical model.

1. For each sample in the posterior in each year, compute the proportion of fish that pass the RST by the end of December. This gives 2,000 x 10 values (one value for each of the 2,000 posterior values for each of 10 years.)
2. Fit a hierarchical model to these values where each yearly mean is sampled from a hyperdistribution with and without covariates for each year of data.

The MT can then estimate multiple proportions by fitting a multivariate hierarchical model to the derived values (perhaps restricting it to be monotonic).

This model can then be extended for the more complex measures of run timing. For example, the median date and steepness can be computed (e.g., difference in days between 25% and 75% percentiles of run timing) similarly to *phi* and *lambda* for each sample of the posterior for each year (this report Figure 3). Now, a simple multi-variate hierarchical model for the bivariate values can be built with the potential for easy addition of covariates.

Neither of the methods above require fitting a beta distribution to describe the entire run distribution in order to estimate proportions passing by certain dates, or percentiles (e.g., median) or steepness. In theory, one could get estimates of the run proportion for every week as a multi-variate (53 element) vector if desired. A spline could be fit to the cumulative run proportion for each posterior sample for each year and then hierarchical models for the spline parameters could be created. Once the hierarchical model is computed, forecasting for a new year is straightforward and is done in much the same way as forecasting for the efficiency trials.

Test more covariates

Some covariates were examined to improve the models but more covariate modeling is needed. Covariates might include additional peak-flow metrics (large migrations can occur during peak flows), river temperature during incubation, fry size early in the season, outmigrant abundance, and the median date of parent spawning.

Explore other covariate types

Expand and refine covariate analyses. Additional evaluation of flow, temperature, and biological predictors (e.g., spawner abundance) is needed, including effects on both median migration timing and the steepness of the outmigration run. The development of forecastable covariates should be prioritized, potentially requiring collaboration with hydrologic modelers.

Explore scenario testing in lieu of covariates

Can year-to-year predictions be made based on case studies of specific years? Interannual variability is high but there could be specific trends in the variability correlated to specific environmental/ecological covariates. For example, do some years that are wet show certain “states” and some that are dry show other “states”? Instead of using flow as a covariate, could a binomial toggle between two potential state spaces in the ecosystem (wet/dry, poor/favorable, etc.) be used for scenario testing in lieu of including an explicit covariate?

Advice for decision makers is needed

Use model forecasts cautiously for early-season decision-making. Early forecasts of annual abundance are highly uncertain and should be applied carefully until management objectives and SRJPE application requirements are clearly defined. This is a challenge, as early-season forecasts are potentially the most useful to end-users but are also usually the worst-performing when fit to observational data.

Guidance for users is needed to address the lower accuracy in some tributaries. How greatly does the low accuracy in several tributaries impact overall decisions? Can the early-season models still be used to inform decision-makers if the best-fit models are not in place until late season?

One specific case where guidance is needed is at RST sites where the out-migration timing model and the abundance model conflict. For example, the Upper Clear Creek model predicted timing of outmigration with good precision during the out-migration portion (Chapter 8 Figure 4), but then the forecasts of abundance are uncertain, and the MT states that no juveniles enter the RST site after February, but the model predicts continuing juvenile entry through March (Chapter 8 Figure 5). These two results must be balanced by end-users so that decision makers understand that the earliest entry tributaries may not see juveniles through the season once their peak entry has passed. Based on abundance alone, decision makers seeking results in mid-February would assume more juveniles are coming through the RST site, but using the timing model would show this is not the case. If the outmigration model and the abundance model provide conflicting information, how should end-users make appropriate decisions about which information to prioritize?

Evaluate alternative approaches for modeling complex migration timing

One issue with forecasting occurs when emigration is bimodal, perhaps in response to multiple peak flows. The simple beta distribution in the current version of the model could not capture more complex outmigration-timing patterns, such as years with both early and late peaks. Exploration of mixture or non-parametric approaches, such as a library of historical outmigration-timing patterns, as noted by the authors, is recommended.

Implement a structured model update process

Models should be routinely refined as new data and covariates become available through a formal review or workshop process.

Complete and refine Feather River modeling components

Further development of the PLAD model is needed before applying the in-season outmigrant model to Feather River data, particularly to address hatchery fish contributions and complex run structure.

Discuss how error propagates forward

How do the incorrect predictions, for example, in Clear Creek, impact the implications both for overall synthesis in this modeling framework and for the use of these independent results in life cycle modeling applications?

Chapter 9 Smolt Survival Model

Chapter Overview

Chapter 9 presents two submodels to estimate Chinook salmon survival from release locations to Sacramento/Delta entry and travel time from release locations to Sacramento/Delta entry. Both submodels are based on the release of acoustic-tagged fish that are then detected at receiver locations downstream.

The data consist of a detection/timing history for each fish, showing the time and release location, and subsequent detection times at four or two receiver locations downstream (depending on where released; refer to MT Chapter 9 Figure 1).

General Comments

Chapter 9, Section 2.1, partially describes the data collected, but additional descriptions are presented in Sections 2.2 and 2.3. These sections should be combined into Section 2.1. It appears that the final data set consists of a single record per fish released at various release locations shown in Chapter 9 Figure 1, for a total of just over 14,000 releases.

The acoustic-tagged fish migrate downstream, where the tag may be detected at various receiver locations. The receiver locations have been grouped into 4 general locations (green dots in Chapter 9 Figure 1). No information has been presented on how to deal with multiple detections at the grouped receiver locations, some of which are far apart, e.g., what date/time is used (earliest, latest), how distance is determined if multiple detections at different receivers in the pooled location occur, etc. In the case of fish taking the Sutter/Yolo Bypass, how is distance defined between receiver locations?

Data summaries are presented on the covariates measured for each release group (Chapter 9 Table 4). However, no summary information is presented on the detection histories, e.g., how many fish had detection histories at the first two stations but then not again until station No. 5. A summary of travel times is also needed in the report.

Various covariates are measured for the fish, e.g., date of release, type of fish, hatchery or wild, location of release, weight, fork-length, and condition factor at the time of release. Were yearling juveniles included in the tag analysis? Chapter 3 indicated that yearling Chinook salmon would be analyzed separately from subyearlings, but the Review Panel understands that the yearling analysis may be prepared at a later date.

Additional environmental covariates are measured, such as peak flow during the month of the release. Does this mean that a fish released at the start of the calendar month and the end of the calendar month have the same value of the flow covariate? Perhaps peak flow during a window centered on the time of release would be a better covariate? This covariate is "… standardized using different means and standard deviations …" (Chapter 9 page 6) but little information on the standardization process is given. For example, is the mean of the peak flows a simple average of the "peak flow" over all years of fish releases in the River/Tributary grouping, or is a mean peak flow based on gauge measurements used and not weighted by number of fish released?

Categorical environmental variables that summarize the flow information are used, such as DWR water year. Is this the "Water Year Type" shown in Chapter 9 Table 1, even though the source of the values in the legend for Chapter 9 Table 1 appears to be different? A "flow exceedance year type" (Chapter 9, page 6) is also used, but this is not presented in any table.

Survival submodel

A common statistical model to estimate the survival of tagged fish as they move downstream past detection locations (receivers) is the Cormack-Jolly-Seber (CJS) model (Lebreton et al. 1992). Because there are multiple release groups across multiple locations and release years, the MT fits a Bayesian Hierarchical model (BHM) version of the CJS. The detection probabilities are modelled as a BHM with random effects for year of release. The survival per 100 kilometers is modeled, initially, as reach-specific with a release-group random effects.

The MT then jumps directly into the BHMs without any justification. Simple analyses that justify the use of a complex Bayesian analysis should be presented. For example, simple non-hierarchical CJS models should be fit for each release group and detection/survival estimates plotted to see the "raw" variability in detection/survival probabilities over time. Detection probabilities for acoustic tags are usually very high (see Chapter 9 Figure 6) so even raw "survival" estimates computed directly from detections would be useful. These values would then show the need for a hierarchical model. Similarly, the simple estimates of survival probabilities should be plotted against the covariates to see if a linear relationship is sensible on the logit-scale etc. Any outliers would also be identified which can be "removed" or "justified".

Both hatchery and wild Chinook salmon are used. Chapter 9 Figure 4 provides some initial justification that the two groups are somewhat similar except for maximum fork-length (see additional comments below). The simple CJS analysis should also be used to examine the simple survival estimates to see if they are similar between wild and hatchery fish.

A model where the survival per 100 kilometers depends on the maximum flow in the month of release and the fork-length of the fish, as fixed effects, and a random effect of the release group-reach combination is found to be the "best" fitting model. This model may "overfit" the data as outlined below.

The product of the reach-specific survival, after adjusting for reach length, from release to Sacramento/Delta entry is found, which will then be used in the prediction model (Chapter 9 Figure 9). Chapter 9 Figure 9 shows that back-casting yields estimates with reasonable uncertainty, but the low variability may be an artifact of the overfitting mentioned above. However, Chapter 9 Figures 10 and 11 show that predictions for future years will be very uncertain with 95% prediction intervals being very wide. This is related to overfitting; the effects of years and release locations are subsumed into the release-group-reach random effects. Better forecasts could be obtained, for example, if year-specific random effects were used, as these have been found to depend on precipitation. Then, predictions using dry, normal, and wet "year" would give a range of survival values that may be more useful for management purposes. Similarly, multiple years of releases occur at several release locations. A random effect for release location may be useful.

The use of a Bayesian CJS model could be useful, but no attempt was made to fit a hierarchical model for the parameters in the survival rates per 100 kilometers) as was done for the detection probabilities.

The standard CJS model is known to have a problem with the non-identifiability of parameters in the last reach and at the last receiver location, i.e., only the product of survival in the final reach and the detection probability at the final receiver location can be estimated. A Bayesian model “resolves” this confounding only because the prior information provides some information to disentangle the two values, and the prior information drives these final estimates. This problem can be seen in Chapter 9 Figure 6 in the very wide posterior distributions for the Delta receiver location and Chapter 9 Figures 7/8 in the very wide posterior distributions for the survival rate per100 kilometers in the Sacramento-DeltaExit reach. It is unclear how “trustworthy” these estimates are, and further analysis is needed to assess their sensitivity to the prior distributions. Perhaps modeling survival rates as a function of covariates resolves this issue, but it needs further analysis. Note that the first paragraph of Chapter 9 Section 3.1 indicates that mixing was an issue when estimating the “through-Delta” migration, which is the Sacramento-DeltaExit reach where the confounding of survival and detection at DeltaExit occurs. Comments on Chapter 9, page 13, about the increased uncertainty in the Sacramento-DeltaExit reach versus the earlier reaches is again an artifact of this confounding. These paragraphs need to be revised.

Travel-time submodel

Estimates of travel time are based on a custom model that uses the detection times of tagged fish. A model where the travel time per100 kilometers depends on fish size covariates and environmental covariates appears to be fit, but no details are provided on which size variables were used or which environmental variables were used in this chapter (see synthesis Chapter 3 page 15 for more details). The same “overfitting” induced by random effects at the release group-reach level may occur. This model gives predicted travel times for each release group (Chapter 9 Figures 12 and 13). The reach-specific travel times are added give an estimate of travel time from release to Sacramento/Delta entry (Chapter 9 Figure 14). No information is presented here on forecasting uncertainty

Similarly, no attempt was made to fit a hierarchical model for the random effects in the travel time models. Consequently, the advantages of using Bayesian models may not be fully realized.

This travel time section also needs careful review by the MT to distinguish between the mean travel time between reaches and the individual travel time between reaches. For example, Chapter 9 Section 3.3 quotes time-to-travel in reaches but presumably these are mean times? But the report then quotes ranges; are these credible intervals or prediction intervals? Similarly, it is often unclear whether posterior distributions for mean travel times or posterior predictive distributions for individual travel times are being used.

A quick analysis exploring the relationship between travel time and survival would be useful. A negative relationship (i.e., shorter travel time and higher survival) when both are related to flow would provide biological insight into the joint processes of travel and survival and suggest which covariates could be useful.

Need for additional modeling effort

Only a small set of models was considered: only one model for detection modeling, Chapter 9 Table 5 for survival modeling, and only one model was considered in the travel time model. Additional models should be considered. For example, is there evidence of a wild/hatchery effect on survival/detections/travel time? Is there evidence of differences in survival/detection/travel time in the common reaches of River and Tributary fish? The justification for combining wild/hatchery fish and River/Tributary fish appears to be based only on plots in Chapter 9 Figures 3 and 4. Only one model explored more than one factor in Chapter 9 Table 5.

The random-effects models in the survival model may be overly flexible. For example, Chapter 9 Equations 2a/2b have a common survival rate in a reach for all years plus random effects for each release group-reach combination. This essentially gives a “perfect” fit to the data because each release group will have its “own” survival rate in a reach in a particular year. However, some release locations have multiple years of data. In many cases, there are location-specific effects (e.g., habitat) and year-specific effects (e.g., warmer than normal), which raise/lower survival rates consistently across reaches and release groups rather than having independent random effects in each release group-reach combination. A model with random effects of release-location and random-effects of release years should also be considered to account for these correlations in survival probabilities across reaches within a year and across releases in the same year. Consider Chapter 9 Figures 5, 7/8. Release groups have additional structure like year, locations, and reaches. The

MT should plot these random effects by year, by locations, and by reaches to see if there are any patterns. This could help determine whether models with structured random effects are worthwhile.

This is one reason why predictions are so uncertain (Chapter 9 Figures 10/11) despite estimates of survival being fairly precise. Sampling from the distribution of the random effects of the release group is essentially non-informative. But year-specific effects may be easier to predict (e.g., may depend on climate conditions such as La Nina, etc.).

Some of the year-specific effects may be subsumed in the flow covariate, but this is not considered. For example, Chapter 9 Figure 10 shows that the uncertainty increases greatly when simple random effects for release groups are added.

No information on the models considered for travel time (similar to Chapter 9 Table 5, for survival rates) is provided. It is therefore unclear which fish size variable and which environmental covariates are used in Chapter 9 Equations 7a/7b. However, the report indicates that “The effects of fish size and flow on travel time were analyzed in greater detail for use in SRJPE, and these results are reported in Chapter 3.” (Chapter 9 page 15), so perhaps this information is presented in Chapter 3. It should also be included in Chapter 9 where more details can be presented. Similar comments about additional models for the random effect are also applicable here.

Fish size issues

The MT relies on acoustic telemetry data to estimate survival across multiple populations and years, and to estimate the travel time of tagged fish from release to Delta entry. Chinook salmon are tracked using Juvenile Salmon Acoustic Telemetry System (JSATS) transmitters (tags; 300 milligrams, 10.7 x 5.0 x 2.8 millimeters) with a 32-day lifespan. The median size of tagged fish was 89 millimeters, and the median tag-to-fish-weight burden was slightly less than 5% (i.e., the burden for an 80-millimeter fish). The MT used data from both spring-run and fall-run smolt releases (including many hatchery fish) to increase the sample size. The MT suggests fall-run juveniles are a good spring-run surrogate because they migrate through the Sacramento River at similar sizes and times of year.

Although the tagged juveniles are representative of the relatively large fish sizes emigrating during the tagging period, mid-March through early June, they are not representative of most juveniles in each spring-run Chinook salmon population. Median migration dates at the tributary RST sites ranged from December 7 to March 15, depending on population (median Delta entry dates ranged from December 21 to April 5; Chapter 3 Figure 8). Median juvenile size in the tributaries during peak emigration ranged from approximately 37 millimeters to 40 millimeters, depending on population. Thus, most of the emigration of spring-run Chinook salmon occurred well before the release of tagged fish, and most emigrating spring Chinook salmon were much smaller than the tagged fish. The modelled results may be appropriate for fish greater than approximately 70 millimeters (how many fish < 80 millimeters were tagged in wet versus dry years?), but extrapolation to smaller fish (e.g., 39 millimeters) may produce biased results.

Although the modeled survival and migration rate relationships with river discharge and body size may generally reflect the pattern for the numerous untagged small juveniles (e.g., higher survival and faster movement during higher flows), we suspect the magnitude of these relationships may be different for the smaller untagged fish. Tagging of large fish occurred late in the season when juveniles were more likely to be actively migrating downstream. In contrast, we suspect most of the smaller, untagged fish that migrated prior to the tagging operation were migrating passively, i.e., carried downstream by the current when suitable rearing habitat was unavailable. In general, we suspect the migration rate of smaller fish is slower because they are unlikely to be actively migrating early in the season. Instead, they are probably seeking river margins to rear and, when possible, forage on prey (e.g., Moyle et al. 2015). It is also possible that the spatial distribution of the smaller Chinook salmon in the river is different from the larger tagged fish; smaller fish may be near the river margins, whereas actively migrating fish may be farther offshore in faster current. Additionally, given potential differences in migratory behavior by size and season, smaller fish may be exposed to a different suite of predators than late-migrating, larger tagged salmon. In summary, the MT should verify the extrapolation of the large, late migrating acoustically-tagged salmon to the much smaller spring-run juveniles that migrate earlier in the season. Please see bullet 4 in the *Recommendations* section below.

Hatchery-origin vs. wild smolts issues

The assumption of equivalence in travel times between hatchery-origin and wild smolts also warrants further investigation. The Review Panel notes that of the 14,520 tagged, approximately 20% are wild and 80% are hatchery-origin (Chapter 9 Table 1). The MT supports the assumption that hatchery fish adequately represented large wild fish with respect to travel time, based solely on similarities in body length during the late-season period. However, long-term telemetry, PIT-tag, and acoustic studies summarized by NOAA Fisheries and regional fish agencies consistently show shorter and less variable outmigration travel times for hatchery juveniles compared to wild conspecifics, after controlling for differences in body length. The Review Panel notes that some hatchery managers watch the behavior of Chinook salmon smolts in the raceway and release them when they become agitated, indicating they are ready to emigrate. The Review Panel recommends explicit acknowledgement that differences in outmigration travel times is very likely to vary between hatchery-origin versus wild smolts, independent of differences in body length, and discussion of how this assumption may invoke undocumented bias into the timing model output and ultimately the SRJPE.

Additional environmental factors impacting travel time

The travel time model accounts for the influence of river discharge on downstream migration timing and survival of salmonid smolts, but could benefit from exploring other environmental factors that may also play a role. For example, research from central British Columbia found that accumulated thermal units, rather than river flow, was the parameter that best explained downstream migration timing for Chinook salmon smolts (Sykes et al. 2009), and work in Oregon found that during a 4-year study on Chinook salmon smolts, higher average spring temperatures were associated with earlier emigration dates (Roper and Scarnecchia 1999). The MT should either incorporate stream temperature, potentially both antecedent and accumulative thermal conditions, into the travel time model, or explicitly state why this addition is not warranted.

Spring-run juvenile salmon may enter off-channel floodplain habitat during high water events. Thus, high water events may delay migration rates of some juveniles, leading to greater growth in prey-rich habitats and potentially greater survival. Key off-channel habitats include Yolo Bypass, Butte Sink, and Sutter Bypass (Tompkins et al. 2017; Cordoleani et al. 2020, https://www.usbr.gov/mp/bdo/yolo-bypass.html). The MT is aware of the interaction between high water events and use of floodplain habitats by juvenile salmon. These events should be further evaluated in the model, if possible, including identification of flow levels that allow juveniles to enter the floodplains.

Potential issues with using fall-run fish as a proxy

Can the MT conduct an analysis to check the assumption that fall-run and spring-run Chinook salmon have similar migration and survival patterns? In much of their range in North America, spring-run populations produce yearling smolts rather than subyearlings. We suspect the subyearling strategy currently observed in the Sacramento Basin is a response to the elimination of the vast majority of spring-run habitat upstream of dams. Survival of spring-run Chinook salmon is being proxied here by fall-run survival. While there is some overlap, fall-run fish do have a different phenology, and conditions can be vastly different in April/May than in Jan/Feb/March. Once they have exited the delta, ocean conditions like food availability, temperature, and predator presence make a big difference in survival and can have high temporal variability.

Recommendations

The Review Panel offers the following comments and recommendations:

1. Additional models for survival and travel time need to be considered.
    a. The “release-group” random effect should be partitioned into year-specific and release-location specific random effects. Examination of the figures in this chapter shows additional structure over a simple release-group effect.
    b. The effect of fish size on migration rate should be evaluated. Chapter 3 Table 2 shows the median migration dates for each population at the RST and the Delta entry. The values suggest very rapid migration from the RST to the Delta, especially for the small

fish that are tagged at the RST (approximately 1 week for Battle Creek and Mill Creek, 0 days for Butte Creek, approximately 2 weeks for Clear Creek). Is rapid travel caused by the reliance on large tagged salmon that are actively rather than passively migrating? What is the travel time for a particle of water travelling from each RST to the Delta entry near the thalweg compared with river margin? We encourage the MT to evaluate potential ways to further evaluate migration rates and survival of smaller Chinook salmon that are more representative of the spring Chinook salmon populations shown in Chapter 3 Figure 8.

c. The effect of hatchery-origin versus wild smolts should be evaluated. Differences in outmigration travel times are very likely to vary between hatchery-origin vs. wild smolts (independent of differences in body length), and this assumption of homogeneity may invoke bias into the timing model output and ultimately the SRJPE.

d. The MT should either incorporate stream temperature, potentially both antecedent and accumulative thermal conditions, into the travel time model, or explicitly state why this addition is not warranted.

e. The MT should conduct an analysis to check the assumption that fall-run and spring-run Chinook salmon have similar migration and survival patterns.

f. The MT should more robustly explore the assumption that low detection is due to a low detection probability and relatively high Delta survival. The issue with low detection rates around the Chipp's receivers in Chapter 9, Section 3.1, and Chapter 9 Figure 6, can be at least partially rectified/checked by using the Chipp's field surveys, which record estimated number of smolts passing the sampling sites on specific dates. A comparison with the survey data from these years would clarify whether they are particularly or anomalously low-ocean-entry years. If yes, the low detection is related to low survival, if no, it is a detection issue instead.

2. Offer guidance on "best" model selection
   The Review Panel recommends that the MT offer guidance on choosing the most useful model; for example, is the best-fit model always the one that provides the most ecological information? If not, how can end-users

develop a decision-making strategy between the “best-fit” model mathematically and the most informative model ecologically?

If a likelihood framework were used, it would be relatively straightforward to use information-theoretic methods, such as AIC, to combine results from multiple models, where each model is given a weight that accounts for complexity and fit. This would also allow fitting more complex models (e.g., multiple covariates) where the increase in fit from more covariates may not overcome the fitting penalty from more covariates and so, more complex models may not be suitable. Multiple model averaging is much more complicated with Bayesian models (e.g., reversible jump MCMC), but simpler approximations are available using Bayesian Model Averaging (BMA, Hoeting et al. 1999).

3. Did the MT consider and mitigate for the potential tagging effect of delayed migration and increased predation by holding the fish for a while until full recovery?
4. Consider alternative approaches to estimate survival and migration rate. Technology continues to improve and could be used to evaluate the MT assumption that the large acoustic-tagged Chinook salmon represent the much smaller spring-run juveniles that emigrate earlier in the season. For example:
    a. Acoustic tag technology continues to be improved, leading to much smaller tags than what JSATS tags currently used in the Sacramento River. Daniel Deng at the Pacific Northwest National Laboratory (zhiqun.deng@pnnl.gov) has developed an experimental small fish "shad tag" that is only 0.05 grams (7.6 millimeters long x 2.0 millimeters in diameter. These tags are being considered for Delta Smelt (*Hypomesus transpacificus*), and laboratory tests have been conducted with smelt as small as 50 millimeters (Deters et al. 2024). The detection range depends on the location, ranging from about 70 meters to 120 meters in freshwater. Tag life at 5-second pulse rate interval is 49 days with Gen-2 microbattery. The existing acoustic tag arrays could probably be used to detect salmon containing these tags.
    b. Passive Integrated transponder (PIT) tags, which can be inserted into salmon as small as approximately 45 millimeters and require no surgery. Key to the use of PIT tags would be the successful deployment of antenna arrays in the Delta entry area. Detection

distances are limited in smaller FDX PIT tags, so sampled fish are often funneled into a constricted area for automatic detection (no fish handling required). Larger HDX tags, which have a somewhat greater detection range, can be deployed in salmon down to about 65 millimeters. Investigators should evaluate PIT tags, which are much less costly than JSATS tags, if they have not already done so. PIT tags have been used to estimate migration timing of wild smolts of the Snake River and other tributaries of the Columbia River (e.g., Achord et al. 1996), including the use of mobile detection arrays (e.g., Saboret et al. 2021).

c. Coded wired tag. The MT mentioned the evaluation of coded wire tagged fish at a meeting with the Review Panel on January 8, 2026, while noting issues with limited CWT recoveries in the Sacramento RST where trap efficiency is low. If sufficient CWT data exist, can it be used to supplement the acoustic tags, especially in areas of low detection? The Review Panel assumes that CWT data may provide travel-time information rather than survival rate data.

d. Otolith microchemistry, daily growth rings, and genetic stock identification might also offer a unique approach that could be further evaluated to estimate residence time in the mainstem river (Ruggerone and Volk 2004). Researchers in the Central Valley and the MT are knowledgeable about these approaches and could evaluate their feasibility (e.g., Sturrock et al. 2019; Cordoleani et al. 2021). The idea would be to sample small juvenile spring Chinook salmon in the Sacramento River at the Delta entry site, then analyze their otoliths for daily growth rings and chemical signatures that could be used to identify tributary origin and, possibly, entry into the Sacramento mainstem, including the days spent in each habitat. Unfortunately, otolith removal would require sacrificing the fish unless it can be determined from adult carcasses. Genetics (or PLAD) might be needed to separate spring from fall Chinook salmon. Otolith microchemistry depends on water chemistry and water temperature. Daily rings could be used to identify days since emergence or days since entry into the mainstem, if a change in microchemistry can be detected on the otolith in response to changes in water chemistry and/or temperature in the tributary, Sacramento River, and estuary. Daily otolith rings also tend to reflect daily growth, and might reflect a shift in habitat from the tributary to the mainstem. Juvenile size at capture could be used to examine growth rate.

References

Achord, S., Matthews, G.M., Johnson, O.W., and Marsh, D.M. (1996). Use of Passive Integrated Transponder (PIT) Tags to Monitor Migration Timing of Snake River Chinook Salmon Smolts. *North American Journal of Fisheries Management* 16:302–313. https://doi.org/10.1577/1548-8675(1996)016%3C0302:UOPITP%3E2.3.CO;2

Bell-Tilcock M, Jeffres, C.A., Rypel, A.L., Willmes, M., Armstrong, R.A., Holden, P., Moyle, P.B., Fangue, N.A., Katz, J.V.E., Sommer, T.R., Conrad, J.L., and Johnson, R.C. (2021). Biogeochemical processes create distinct isotopic fingerprints to track floodplain rearing of juvenile salmon. *PLOS ONE* 16(10): e0257444. https://doi.org/10.1371/journal.pone.0257444

Bonner, S.J., and Schwarz, C.J. (2011). Smoothing population size estimates for Time-Stratified Mark-Recapture experiments Using Bayesian P-Splines. *Biometrics* 67: 1498-1507. https://doi.org/10.1111/j.1541-0420.2011.01599.x

Brandes, P., Pyper, B., Banks, M., Jacobson, D., Garrison, T., and Cramer, S. (2021). Comparison of Length-at-Date Criteria and Genetic Run Assignments for Juvenile Chinook Salmon Caught at Sacramento and Chipps Island in the Sacramento–San Joaquin Delta of California. *San Francisco Estuary and Watershed Science*,19 (3). https://doi.org/10.15447/sfews.2021v19iss3art2

Brooks, S. P., and Gelman, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, 7(4), 434–455. https://doi.org/10.1080/10618600.1998.10474787

Chen, E.K., Lumahan, K., Johnson, R.C., Phillis, C.C., Whitman, G., Sturrock, A.M., Satterthwaite, W.H., and Carlson, S.M. (2025). Juvenile life history, migration, and habitat use of natural-versus hatchery-origin Chinook Salmon. *Transactions of the American Fisheries Society*, Volume 154, Issue 4, July 2025, Pages 440–455, https://doi.org/10.1093/tafafs/vnaf021

Cordoleani, F., C.C. Phillis, A.M. Sturrock, A.M. FitzGerald, A. Malkassian, G.E. Whitman, P.K. Weber, and R.C. Johnson. (2021). Threatened salmon rely on a rare life history strategy in a warming landscape. *Nature Climate Change*. 11, 982–988. https://doi.org/10.1038/s41558-021-01186-4

Cordoleani, F., Satterthwaite, W.H., Daniels, M.E., and Johnson, M.R. (2020). Using Life-Cycle Models to Identify Monitoring Gaps for Central Valley Spring-Run Chinook Salmon. *San Francisco Estuary and Watershed Science* Volume 18, Issue 4. https://doi.org/10.15447/sfews.2020v18iss4art3

Deters, K.A., Janak, J.M., Mueller, R.P., Boehnke, B.T., and Deng, Z.D. (2024). Handling and Tagging Techniques for Implanting Juvenile American Shad with a New Acoustic Microtransmitter. *Journal of Visualized Experiments*. 208:e65694, https://doi.org/10.3791/65694

Gelman, A., and Rubin, D.B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* 7(4):457–511. https://sites.stat.columbia.edu/gelman/research/published/itsim.pdf

Gimenez, O., Morgan, B.J.T., and Brooks, S.P. (2009). Weak identifiability in models for mark-recapture-recovery data. In *Modeling demographic processes in marked populations Environmental and Ecological Statistics*, vol 3. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-78151-8_48

Harvey, BN, Jacobson, D.P. and Banks, M.A. (2014). Quantifying the uncertainty of a juvenile Chinook Salmon race identification method for a mixed-race stock. *North American Journal of Fisheries Management* 34(6):1177–1186. https://doi.org/10.1080/02755947.2014.951804

Harvey, B, Nelson, P., Gill, S., Vizek, A., and Cain, E. (2022). Data Management Strategy for the Spring-Run Chinook Salmon Juvenile Production Estimate. October. water.ca.gov/-/media/DWR-Website/Web-Pages/Programs/State-Water-Project/Endangered-Species-Protection/ITP-SR-JPE-Data-Mgmt-Strategy-2022-09-07.pdf.

Harvey, B. and Stroble, C. (2013). Comparison of genetic versus Delta Model Length-at-Date race assignments for juvenile Chinook Salmon at state and federal south Delta salvage facilities. CDFW Report. Accessed at: https://www.waterboards.ca.gov/waterrights/water_issues/programs/bay_delta/california_waterfix/exhibits/docs/PCFFA&IGFR/part2/pcffa_136.pdf

Hoeting, J. A., Madigan, D., Raftery, A.E., and Volinsky, C.T. (1999). Bayesian Model Averaging: A Tutorial. *Statistical Science*, *14*(4), 382–401. http://www.jstor.org/stable/2676803

Jeffres, C.A., Opperman, J.J., and Moyle, P.B. (2008). Ephemeral floodplain habitats provide best growth conditions for juvenile Chinook salmon in a California river. *Environmental Biology of Fishes* (2008) 83:449–458. https://doi.org/10.1007/s10641-008-9367-1

Kururvilla, M., Quinn, T.P., Anderson, J.H., Miller, E.M., Berger, A.G., Okasaki, C., McMillan, J.R., Pess, G.R., Westley, P.A.H., and Berdahl, A.M. (2024). Social influences complement environmental cues to stimulate migrating juvenile salmon. *bioRxiv* 2024.06.14.598894; doi: https://doi.org/10.1101/2024.06.14.598894

Lacoursiere-Roussel, A., Cote, G., Leclerc, V., and Bernatchez, L. (2016). Quantifying relative fish abundance with eDNA: a promising tool for fisheries management. *Journal of Applied Ecology,* 53, 1148–1157. https://doi.org/10.1111/1365-2664.12598

Lebreton, J.D., Burnham, K.P., Clobert, J., and Anderson, D.R. (1992). Modeling Survival and Testing Biological Hypotheses Using Marked Animals: A Unified Approach with Case Studies. *Ecological Monographs*, *62*(1), 67–118. https://doi.org/10.2307/2937171

Limm, M.P., and Marchetti, M.P. (2009). Juvenile Chinook salmon (*Oncorhynchus tshawytscha*) growth in off-channel and main-channel habitats on the Sacramento River, CA using otolith increment widths. *Environmental Biology of Fishes* 85: 141-151. https://doi.org/10.1007/s10641-009-9473-8

Ludwig, D., and Walters, C.J. (1981). Measurement Errors and Uncertainty in Parameter Estimates for Stock and Recruitment. *Canadian Journal of Fisheries and Aquatic Sciences*. 38(6): 711-720. https://doi.org/10.1139/f81-094

Monzyk, F.R. Jonasson, B.C., Hoffnagle, T.L., Keniry, P.J., Carmichael, R.W., Cleary, P.J. (2009). Migration Characteristics of Hatchery and Natural Spring Chinook Salmon Smolts from the Grande Ronde River Basin, Oregon, to Lower Granite Dam on the Snake River. *Transactions of the American Fisheries Society*, Volume 138, Issue 5. https://doi.org/10.1577/T08-108.1

Moyle, P.B., Quiñones, R.M., Katz, J.V.E. Katz, and Weaver, J. (2015). Fish Species of Special Concern in California. 3rd edition. Sacramento: California Department of Fish and Wildlife. https://www.wildlife.ca.gov/Conservation/Fishes/Special-Concern

Nakagawa, S., and Schielzeth, H. (2012). A general and simple method for obtaining $R2$ from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, in press. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Nelson, P., Baerwald, M., Burgess, O., Bush, E., Collins, A., Cordoleani, F., DeBey, H., Gille, D., Goertler, P., Harvey, B., Johnson, R., Kindopp, J., Meyers, E., Notch, J., Phillis, C., Singer, G., and Sommer, T. (2022). Considerations for the Development of a Juvenile Production Estimate for Central Valley Spring-Run Chinook Salmon. *San Francisco Estuary and Watershed Science* 20. https://doi.org/10.15447/sfews.2022v20iss2art2

Olsen, J.B., Miller, S.J., Harper, K., Nagler, J.J., and Wenburg, J.K. (2006). Contrasting sex ratios in juvenile and adult chinook salmon *Oncorhynchus tshawytscha* (Walbaum) from southwest Alaska: sex reversal or differential survival? *Journal of Fish Biology* 69: 140–144. http://dx.doi.org/10.1111/j.1095-8649.2006.01119.x

Phillis, C.C., Sturrock, A.M., Johnson, R.C., and Weber, P.K. (2018). Endangered winter-run Chinook salmon rely on diverse rearing habitats in a highly altered landscape. *Biological Conservation* 217: 358-362. https://doi.org/10.1016/j.biocon.2017.10.023

Roper, B., and Scarnecchia, D. (1999). Emigration of age-0 chinook salmon (Oncorhynchus tshawytscha) smolts from the upper South Umpqua River basin, Oregon, U.S.A. *Canadian Journal of Fisheries and Aquatic Sciences* 56. 939-946. https://doi.org/10.1139/f99-023

Rose, K., Anderson, J., McClure, M., and Ruggerone, G. (2011). Salmonid integrated life cycle models workshop. Report of the Independent Panel. Independent scientific review prepared for the Delta Science Program, CA. https://www.waterboards.ca.gov/bay_delta/docs/cmnt091412/sldmwa/rose_et_al_2011.pdf

Ruggerone, G.T., and Volk, E.C. (2004). Residence time and growth of natural and hatchery Chinook salmon in the Duwamish Estuary and Elliott Bay, Washington: an application of otolith chemical and structural attributes. Prepared for U.S. Army Corps of Engineers, Seattle District, and Port of Seattle. Prepared by Natural Resources Consultants, Inc. and Washington Dept. Fish and Wildlife. Seattle, WA. https://www.researchgate.net/publication/261026268_Residence_time_and_growth_of_natural_and_hatchery_Chinook_salmon_in_the_Duwamish_Estuary_and_Elliott_Bay_Washington_based_on_otolith_chemical_and_structural_attributes

Saboret, G., Dermond, P., and Brodersen, J. (2021). Using PIT-tags and portable antennas for quantification of fish movement and survival in streams under different environmental conditions. *Journal of Fish Biology* 99:581–595. https://doi.org/10.1111/jfb.14747

Seber, G.A.F. (1982). *The Estimation of Animal Abundance and Related Parameters*. 2nd ed. London: Griffin.

Sommer, T.R., Nobriga, M.L., Harrell, W.C., Batham, W., and Kimmerer, W.J. (2001). Floodplain rearing of juvenile Chinook salmon: evidence of enhanced growth and survival. *Canadian Journal of Fisheries and Aquatic Sciences* 58:325–333 https://doi.org/10.1139/f00-245

Spear, M.J., Embke, H.S., Krysan, P.J., and Vander Zanden, M.J. (2021). Application of eDNA as a tool for assessing fish population abundance. *Environmental DNA.* 3:83–91. https://doi.org/10.1002/edn3.94

Sturrock, A.M., Carlson, S.M., Wikert, J.D., Heyne, T., Nusslé, S., Merz, J.E., Sturrock, H.J.W., and Johnson, R.C. (2019). Unnatural selection of salmon life histories in a modified riverscape. *Global Change Biology*. 26:1235–1247. https://doi.org/10.1111/gcb.14896

Sturrock AM, Wikert, J.D., Heyne, T., Mesick, C., Hubbard, A.E., Hinkelman, TM, Weber, P.K., Whitman, G.E., Glessner, J.J., and Johnson, R.C. (2015). Reconstructing the migratory behavior and long-term survivorship of juvenile Chinook salmon under contrasting hydrologic regimes. *PLoS One* 10(5): e0122380. https://doi.org/10.1371/journal.pone.0122380

Sykes, G.E., Johnson, C.J., and Shrimpton, J.M. (2009). Temperature and flow effects on migration timing of Chinook Salmon smolts. *Transactions of the American Fisheries Society* 138:1252–1265. https://doi.org/10.1577/T08-180.1

Tompkins, M., Anderson, J., Goodwin, P., Ruggerone, G., Speir, C., and Viers, J. (2017). Yolo bypass salmon habitat restoration and fish passage analytical tool review. Convened by the Delta Science Program. https://cawaterlibrary.net/wp-content/uploads/2018/03/FINAL_REPORT_10-20-17.pdf

Tsuji, S., Inui, R., Nakao, R., Miyazono, S., Saito, M., Kono, T., and Akamatsu, Y. (2022). Quantitative environmental DNA metabarcoding shows high potential as a novel approach to quantitatively assess fish community. *Scientific Reports* 12, 21524 (2022). https://doi.org/10.1038/s41598-022-25274-3

Vinarcsik, J. H., Jaenecke, K.E., Palik, A.F., Sandford, B.P., Bentley, P.J., Brooks, G., and Morris, M.S. (2025). Detection of PIT-Tagged Juvenile Salmonids Migrating in the Columbia River Estuary, 2024. U.S. Department of Commerce, NOAA Contract Report NMFS-NWFSC-CR-2025-02. https://doi.org/10.25923/tkz0-rx68

Woodson, L.E., Wells, B.K., Weber, P.K., MacFarlane, R.B., Whitman, G.E., and Johnson, R.C. (2013). Size, growth, and origin-dependent mortality of juvenile Chinook salmon *Oncorhynchus tshawytscha* during early ocean residence. *Marine Ecology Progress Series* 487:163-175. https://doi.org/10.3354/meps10353

Appendices

Appendix A -Chapter-specific minor comments and editorial suggestions

Chapter 3 Editorial and Minor Comments

- References to Korman et al. (2025a, 2025b, 2025c) and others are not in the Reference section. Presumably, these refer to the different chapters of the Modeling Team Draft Report (2025)?
- “Survival rate” should be replaced by “survival probability” or simply “survival” because these are not per year or per kilometer.
- Text needs to be carefully edited to maintain consistent terminology and distinguish between outmigration (all runs) and spring-run outmigration. Similarly, the report uses ”abundance” and “outmigration size” interchangeably. Include “spring-run” whenever this is appropriate rather than through inference. For example, in Chapter 3 Figure 14, the MT uses the term “spring-run annual juvenile abundance”; “annual outmigration abundance”. In Chapter 3 Figure 15, the MT uses “annual outmigrant abundance”, “juvenile outmigrant abundance”, “outmigrants”, etc., which are unqualified about being “spring-run”. Term switching makes it hard for the reader to keep on top of things.
- Analysis repeatedly references the date at which 50% of spring-run outmigrants have arrived. 10% and 90% arrival dates would also be useful to present.
- p. iii “The spring-run juvenile abundance estimates at the Yuba River site are likely substantially overestimated and were excluded from further analysis in this chapter.” Is PLAD the cause for this?
- p. iii “To increase confidence in JPE forecasts, predictions from SRJPE and its submodels should be tested using out-of-sample data. These tests may

reveal model limitations and ways to improve predictions." The Review Panel agrees completely.

- p. 4 "The in-season submodel forecasts cumulative outmigration proportions for each week of the outmigration period in a forecast year." There appear to be only four forecast time points, which reduces the work in the submodel. As noted in other comments, it is not necessary to fit the beta distribution, and the posteriors for the proportions at the four forecast times can be found directly.
- Chapter 3 Equation 5 is not a data likelihood. At this point, the models compute a posterior distribution for the abundance up to time *iwk* from the BT-SPAS-X/PLAD model. The authors summarize this using a normal distribution, but there is no need to do this – samples from the BT-SPAS-X/PLAD models can be used directly without summarization. There is NO fitting to data involved here.
- p. 6 "Note here the independent estimate of annual abundance is a derived variable and thus does not have a prior distribution. As a result, random samples from the distributions in the numerator and denominator of Chapter 3 Equation 6 will have identical distributions to their priors." This is a bit misleading. There is no direct prior for the independent estimate of annual abundance, but priors on parameters used by BT-SPAS-X/PLAD are incorporated. The second sentence about "random samples" is meaningless and should be struck.
- p. 9 "Travel time decreases …". Here and elsewhere should read "Median travel time decreases …".
- p. 9 "The latter model uses an annual forecast of peak monthly flows to predict travel time from RST locations to Delta entry …" Here and elsewhere in the chapter, caution is needed to differentiate median travel time (Chapter 3 Figure 11) and individual travel time.
- p. 10 "… of juveniles …". Add qualifier "of spring-run juveniles" here and elsewhere.
- p. 10 "The week of arrival …" should read "The median week of arrival …" here and elsewhere.

- p. 10 "… so it can be used to identify an array element." What array?
- p. 10 Chapter 3 Equation 10. "Note that TT is a posterior distribution". Posterior distributions refer to uncertainty about parameters. Perhaps the posterior predictive distribution is meant here, referring to the distribution of individual travel times?
- p. 11 Chapter 3 Equation 12. Isn't the denominator in the first equation equal to 1?
- p. 11 Chapter 3 Equation 12. The notation is confusing because the index and range of summation are in the variable indices rather than on the summation sign, which differs from Chapter 3 Equation 13 (top equation), which then switches notation in the second equation.
- p. 11 Chapter 3 Equation 13. The weighting factor cannot be the raw tributary abundances but need to be normalized to sum to 1 otherwise Chapter 3 Equation 13 (top) doesn't make sense.
- Chapter 3 Figure 5. How does the estimate of total spring-run outmigration at Delta entry compare to the abundances computed on the mainstem RST?
- Chapter 3 Figure 7. In the caption of the figure, it would be helpful to explicitly identify that "Peak Flow" for the Battle/Clear/Mill/Deer Creek group refers to flow in the Sacramento River downstream. Identify the location on the Sacramento River where the peak flow was determined.
- Chapter 3 Figure 9. The uncertainty in survival estimates from the RST to Delta entry were exceptionally high, e.g., median survival for Battle, Clear, and Mill creeks were approximately 2.5% but the upper 80% CI was approximately 23%. Additionally, these survival rate values stemmed from tagged fish that were much larger (median 89 millimeters) than most migrants (approximately 40 millimeters). Thus, uncertainty in survival (and travel times) is likely greater than that reported here.
- Chapter 3 Figure 11. Are the shaded areas 80% credible intervals for the median or 80% prediction intervals for individual travel times? This is similar to the distinction between a confidence interval for the mean response in a regression, versus the prediction interval for an individual

response. These appear to be individual predictions, so the shaded areas are prediction intervals rather than credible intervals.

- Chapter 3 Figure 11. Interesting that travel time increases with higher flows in Butte and Yuba River, but the MT suggest this is "... likely the result of fish routing through sloughs and bypasses which are only accessible at higher flows". If this is true, then why did the other populations not experience the same delay during high flows? Also, the figures should not extrapolate below the size of fish from which travel time was estimated (approximately 70 millimeters), and especially not below the smallest potential size of spring-run Chinook salmon (i.e., X-axis down to 20 millimeters which is not realistic).
- Chapter 3 Figure 12. Same comments as in Figure 11.
- Chapter 3 Figure 15. The four graphs in each row should have a common X-axis so that they are more easily compared across time.
- Chapter 3 Figure 18. The MT discusses multiple potential reasons for the discrepancy in migration timing and suggests that a more plausible bias is a decline in catch efficiency at the mainstem RST later in the migration season as fish become larger and are more able to avoid the RST. The MT should review the literature to evaluate the size of smolts that begin to avoid the RST traps used in the (e.g., Monzyk et al. 2009). Trap avoidance also increases when river velocity is lower.

Chapter 4 Editorial and Minor Comments

- p. 2 Chapter 4 Equation 1. U and p should have a circumflex to indicate that these are estimates.
- p. 3 Chapter 4 Equation 2. U, U_t, p_t need circumflexes.
- p. 3 Throughout Chapter 4 Section 2.2, equations are presented without accompanying variable explanations, or with variable explanations that occur much later in the associated text. Follow the example of Chapter 4 Equation 8 or the sister model explanations in Chapter 05.

- p. 5. Use circumflexes to indicate estimates to distinguish from the actual value of the parameter.
- p. 6 "HBMs jointly estimate the posterior and prior distributions". Bayesian methods do NOT estimate prior distributions. They may estimate the parameters in hierarchical models, but at the end the priors MUST be specified.
- p. 6 "Using the RST capture probability example from above, the prior distribution represents the extent of variation in capture probability across the 10 weekly trials, while the data likelihood represents the fit of 10 different capture probabilities given the 10 observations of r and R." The hierarchical distributions are not a PRIOR distribution. As the MT notes later on, specific prior distributions for the parameters of the hyper distribution must still be specified.
- p. 7 Chapter 4 Equation 4. In this chapter the MT uses "norm()" for a normal distribution. In other chapters, the MT (incorrectly) uses dnorm(). First, be consistent across chapters. Second, use proper distribution names such a Normal() or Binomial() etc.
- p. 7 Just following Chapter 4 Equation 5 is an extra line break.
- p. 7 "For a given number of coin tosses, certainty in the estimate of p_t will increase with the true value of p_t." The variance of an estimated proportion is $p(1-p)/n$ and is maximized when $p = 0.5$. The uncertainty of the estimated p when $p = .05$ is much lower.
- p. 9 Chapter 4 Equation 7. Back to dbin() etc. See earlier comments.
- p. 10 "y = a + bx + cx^2 + d*X^3". Use proper math notation.
- p. 19 "1E06". Convert to a proper number.
- p. 26 The computation of r^2 should be done on the logit() scale rather than the [0.1] scale. The MT notes that r^2 are close to 100%. The reported values of r^2 are misleading because the random effects are essentially chosen to get an exact fit (as noted earlier in the chapter). Alternate measures of fit such as r^2 in the presence of random effects similar to those presented in Nakagawa and Schielzeth (2012) should be used to give a more realistic measure of fit.

- p. 32 Reference to Chapter 8 should be to Chapter 5?
- Chapter 4 Figure 2. Labels for the red contours are cut off.
- Chapter 4 Figure 4. Figure labelling seems off between Figures 4 and 5.
- Chapter 4 Figure 5. Y axis should be on logit scale.
- Chapter 4 Figure 6. Y axis should be on logit scale.
- Chapter 4 Figure 7. Y axis should be on logit scale. Use violin plots to display posterior distributions compactly here and elsewhere.
- Chapter 4 Figure 8. Draw the fitted spline. Panel labels don't line up with the text references for Figure 8.
- Chapter 4 Figure 9 and Figure 11. Plot on the log() scale. Use geometric means to summarize. Make axes common across sites to examine, if for example, certain years tend to have higher runs at all sites?

Chapter 5 Editorial and Minor Comments

- p. 1 The second paragraph in the introduction contains a small grammar error: "This model estimates the weekly abundance of juvenile abundance …" This should likely read "This model estimates the weekly abundance of juvenile salmon …" or should read "This model estimates the weekly juvenile abundance …".
- p. 2 (and other places). Chapter 5 Equation 1b uses dnorm(x,y). This is non-standard. The proper distribution name should be used and identify the parameters so there is no confusion. For example, avoid N(0,16) because is the 16 a standard deviation, a variance, or a precision (1/variance)? Specify directly as Normal (μ = 0, σ = 4). Similar comments for dbin, dgamma, dflat. An additional note is that that dflat() is NOT a valid probability distribution. This is just a "shortcut" for saying an improper non-informative prior.
- p. 3 Q is a standardized flow at site with a separate standardization for each site. How was flow determined for each site.week. Is this from gauge data? From measurements at the RST? How were missing values handled?

- p. 3 Chapter 5 Equation 2a. The prior for the mean (on the logit) scale at $Q = 0$ will place much weight at values of p close of 0 and to 1 rather than uniformly between (0,1). The variance of the prior distribution for this term should be substantially smaller.
- p. 5 Chapter 5 Equation 4. Only the flow has an explicit s subscript because it is subsumed into the t subscript. Be consistent.
- p. 6 "The log of unmarked abundance is estimated in units of 1,000 for numerical precision …". Changing to units of '000 simply changes the log(abundance) values by a constant of $log(1000) = 6.91$ and has no impact on numerical precision. Are there numerical issues that are plaguing BT-SPAS-X?
- p. 7 Chapter 5 Equation 6 should have a mean of 0.
- p. 8 Chapter 5 Equations 8 and 9. The p_t should be π_t.
- p. 9 "The posterior distributions of these parameters are passed to an R script to calculate capture probability for each week of the site-run year being modeled." This chapter and the previous chapter refer to "the RScript" that bridges the stan estimates of capture probability from the HBM model and the BUGS spline model used for abundance estimation. This feels out of place, as every portion of the modeling framework has been meticulously documented and named, and now we have an un-named R-script with little detail just called "the R-script." It is even referred to as such in the acyclic figure. Since the entire BT-SPAS-X model is called using R (even the stan and BUGS components are called using R libraries), a different name or more detail is needed here than just "an R-script".
- p. 11 The computation of r^2 should be done on the logit() scale rather than the [0.1] scale. The MT notes that "The good fit occurred because the capture probability model estimates random-effect deviates for each efficiency trial." This implies that the r^2 are misleading because the random effects are essentially chosen to get an exact fit. Alternate measures of fit such as r^2 in the presence of random effects similar to those presented in Nakagawa and Schielzeth (2012) should be used to give a more realistic measure of fit.

- p. 13 The computation of r^2 between abundances should be done on the log() scale.
- Chapter 5 Tables 2a/2b. Express the efficiency value to a common set of decimal places, i.e., sometimes 2 decimals, sometimes 1 decimal or 0 decimal places are reported.
- Chapter 5 Tables 2a/2b. Some of the efficiency values appear to be odd. For example, consider the value at the top of page 4. There were 2 recoveries/1,106 releases so the efficiency value is 2/1,106 = 0.18%, but the value is reported as zero. There appear to be many cases of this occurring in this table.
- Chapter 5 Figure 3 and other figures., All figures need labels for panels (a, b, c, etc.). In chapter 4, plots were referenced as (a) and (b) etc., based on site, but here they are referenced by panel number even if they are the same site. Consistent referencing and labelling are needed among and within chapters.
- Chapter 5 Figure 3. Labels for the precision of capture probability contours are cut-off (similarly, they are also cut off in the comparable figure in Chapter 04). Please include the full label.
- Chapter 5 Figure 4. Because the analysis is done on the logit-scale, a similar plot should.be produced on the logit-scale for the Y-axis.
- Chapter 5 Figure 5 Appendix A. Present a similar log on the log(abundance) scale. Show the fitted spline curve. None of the values in Appendix A show the spline curve as well.
- Chapter 5 Figures 6 and 8. Align the X-axis by year. This way it is easy to see that, for example, both sites had large run sizes in 2014. The geometric mean abundance across all years should be used rather than the simple mean to avoid extreme influences from extreme run-sizes.
- Chapter 5 Figure 7. There appears to be a definite smooth progression in the spring-run proportion over weeks in a year, but the uncertainty appears to be very large in each week. See comments in the PLAD model about this.

- Chapter 5 Figure 8. (and elsewhere). The report needs to be careful in describing the results. For example, the legend in this figure indicates that the "The bar height and error bars represent the means and 95% credible intervals." The "mean" is the mean of the posterior distribution. This qualification should be added throughout the multiple chapters.

Chapter 6 Editorial and Minor Comments

- p. 7 Chapter 6 Equation 6. Center the values of t by subtracting the middle of the year, e.g., $t - 25$ and $(t - 25)^2$. This will reduce correlation among the parameter estimates and speed convergence.
- Chapter 6, Section 4.3. The text references that Chapter 6 Equation 8 is similar to Chapter 6 Equation 4, but it appears to be referencing Chapter 6 Equation 3.
- p. 12 Chapter 6, Section 4.4. Additional model assessments needed such as traceplots, posterior predictive plots (Bayesian p-value), prior-posterior overlap, etc. A problem with mixture models is inadvertent label switching. Did this occur?
- Chapter 6 Table 2 shows that some parameters don't mix very well across the chains, i.e. small effective sample size. Perhaps longer chains with more extreme thinning will be required.
- p. 2 Chapter 6 Equation 1. Text describes the p as "The probabilities, p_t of run type are calculated as the probability of the observed fork length for individual j given the distribution of fork lengths for each run type." The p's are the probability of belonging to category i as described previously on page 1. This text is contradictory.
- p. 4 Chapter 6 Equation 3. The $FL_{t,i}$ needs a j to index individual j classified as group i at t to create a proper data likelihood like Chapter 6 Equation 8. Chapter 6 Chapter 6 Equation 3 can likely be dropped.
- p. 5 $\pi_{,t,i}$ should likely read $\pi_{t,i}$.
- p. 6, BI should read B_i.

- p. 6 "… assumes that the variance in fork lengths" should read the "… variance on the log scale …". This is not an unreasonable assumption.
- p. 7 Chapter 6 Equation 5. The ρ's change a 3-proportion problem to 2 parameters. Give the equations for the reader showing the reverse.
- p. 8 Introduction to Bayesian analysis. Many submodels use Bayesian analysis and should have this info in their chapter. Perhaps pull this section from this chapter and put it into Chapter 1 or 2.
- p. 10 "… influenced by both the prior and the posterior …". This should read "*prior*" and "*likelihood*".
- p. 11 "… 95% posterior confidence interval is 0.73 to 0.92." Should read "credible interval" rather than confidence interval.
- Chapter 6 Table 5. Report values to a consistent number of decimal places, e.g., 0.20 rather than 0.2.
- Chapter 6 Figure 9. Fix label on legend for Run ID. Don't plot by Julian date, because that splits the run into two segments. Plot by days into the brood or water year like Chapter 6 Figure 10.
- Chapter 6 Figure 10. Extend the tick marks on tick mark label values so it is clear which tick mark is associated with a value. For example, which tick mark refers to Julian week 50. See other plots for ways this can be done.

Chapter 7 Editorial and Minor Comments

- Korman et al. (2025c) is not in the References section.
- p. ii Why is only one mainstem site being used?
- p. iii How do the results compare when the index of spawning abundance is used compared to the full spawning abundance in the tributaries and predictions are made from both analyses?
- p. 2 Chapter 7 Equation 1/ Chapter 7 Equation 2a. What happens if the covariate is strongly predictive of spawning abundance. Then it may be difficult to detect the impact of spawning abundance if the covariate is

included in the model. Figures should include spawner numbers to see if this is true.

- p. 4 Chapter 7 Equation 2d. Why is the “discrete” Ricker model treated differently from the covariate model? Including discrete covariates is standard and requires only a set of indicator variables and coding to map them to the levels of the discrete factor. This is quite common in regression models.
- p. 4 Chapter 7 Equation 3. The random variation is in the estimated R/S and not the “true” R/S, i.e., the normal distribution is around the “True” R/S and not the estimated R/S.
- p. 5 “… was not available because upstream passage estimates in Butte, Mill, and Deer creeks were not considered reliable by regional monitoring teams. Thus, we indexed spawner abundance contributing to juvenile production at Knights Landing based on the sum of annual upstream passage estimates from Battle and Clear creeks only.” Explain what “not considered reliable” means – large bias? large uncertainty? If the latter, then a model with errors in estimates of spawners could be used.
- p. 7 “… were used for all parameters except σ_p where a uniform prior with a lower limit of 0.01 was used." Was the upper limit set to 1? Why was this chosen?
- P. 7 The text says that there are 6 steps in the LOOCV approach to calculating sample error, then only 5 steps are listed.
- p. 8 Bullet 4. Compute the difference on the log() scale. This will automatically scale vastly different abundances.
- P. 9 There may be overfitting for the Yuba site based on the low amount of data available (n = 7).
- p. 15 “One explanation here is that spawner abundance during the period the model was fit to was far from carrying capacity, resulting in negligible density- dependent effects on egg-outmigrant survival rates.” What do people knowledgeable about this population say about this hypothesis? Is some measure of spawning habitat availability/usage available?

- p. 17 "State-space approaches to fit stock-recruit models have been developed to account for uncertainty in stock size (spawner) estimates to avoid bias and underestimation of uncertainty in parameters (e.g., Staton et al. 2017). We could not employ this approach in our analysis because estimates of uncertainty in annual spawner abundance estimates were not available." As noted above, try different levels of assumed error in the spawner numbers.
- p. 19 In a check on the reasonableness of calculated outmigrants per spawner the text on page 19 predicts egg to fry survival assuming 2,752 eggs per spawner. This fecundity value seems too low for female Chinook salmon, given that an age-4 female (3 winters at sea) produces about 6,000 eggs and age-3 females (2 winters at sea) produce 4,000 eggs, according to the Stillwater report. The Executive Summary used a fecundity of 3,000 eggs per female in Battle Creek. Please clarify whether fecundity is per female, or per male plus female spawner. Also, consider that male and female spawners are not always equal, which can lead to another potential source of error in the spawner counts used in the model.
- Chapter 7 Tables 3a etc. Rows may not have been consistently colored with yellow. Legend says that "Yellow-shaded rows identify models that have the lowest MAE_rel values within a site-spawner data type set." What does this mean? For example, in Chapter 7 Table 3a, models with *MAE_REL* of 46% (*si_above_13_temp_day*) are not highlighted but these values of *MAE_REL* are comparable to the values in yellow.
- Chapter 7 Table 11. Add meaning of *rd/us*, etc., to the table similar to site names at bottom.
- Chapter 7 Figure 2. Include spawner numbers in the plot to see if the covariates and spawner numbers are correlated.
- Chapter 7 Figure 3. There are two figures labelled Figure 2. Perhaps Figure 2a is Figure 3a, etc.
- Chapter 7 Figure 3 is mislabeled. The titles of the plots (3a-3g) don't line up with the creeks they are said to correspond to in the text, and plot 3h is missing in the plots section but is referred to several times in the results. It appears that a plot is missing between 3c and 3d, as the plot labelling

seems to be one off from the point at which this is mentioned in the text. 3d is for Deer Creek, not Clear Creek. 3e is for Butte Creek but is referenced in the text as Deer, and Knights Landing is plot 3g but referenced in the text as 3h.

- Chapter 7 Figure 5. This appears to be the Bayesian equivalent of "sampling correlations" among regression parameters. Of real interest would be correlations in the posterior between the coefficients for Spawners in the model and the covariate. A high correlation here would indicate that these are nearly non-identifiable. The correlation between these two parameters should be presented for sites.
- Chapter 7 Figure 6. It seems counterintuitive that when a covariate is added, the *RMAE* increases (1999, 2004, 2008). Further explanation is needed - the explanation on page 14 doesn't match the figure.
- Appendix. The online appendix for covariate dataset collation lists time of emergence as a possible covariate but this isn't used in the models; is this still an option? Because only one environmental covariate is fit per model and this particular covariate combines information on temperature and its impact on developmental effects, using this covariate would essentially allow for an implicit interaction effect without needing to have two covariates.

Chapter 8 Editorial and Minor Comments

- Korman et al. (2024 a, 2024b) is not in the References section.
- p. 3 Chapter 8 Equation 1. Define what is a beta(x, alpha, beta). Is this the cumulative distribution function? Why is this summed over all 53 weeks? The equation needs to be reviewed and better explained.
- p. 4 Chapter 8 Equations 3a/3b. Hmmm… Let a beta distribution be parameterized by B(alpha,beta) = Gamma(alpha+beta)/Gamma(alpha)/Gamma(beta) $x^{(alpha-1)}(1-x)^{(beta-1)}$. This has a mean of alpha/(alpha+beta).
- So if *phi* = mean of the distribution, *phi* = alpha/(alpha+beta). This implies that *lambda =* (alpha+beta) which the authors define as the "sample size".

Refer to the Wikipedia article at https://en.wikipedia.org/wiki/Beta_distribution.

- From the same Wikipedia page, there is NO simple form for the median, but if alpha,beta>1, the median is approximated by (alpha-1/3)/(alpha+beta-2/3) or (for large alpha and beta), will be close to the mean.
- The authors continue
- "Phi represents the proportional week of the year when 50% of the outmigrant population has passed the RST site. This is often termed the mean of the beta distribution, and in the context of the outmigration model represents the median outmigration date."
- The above interpretation is not quite correct, phi only approximates the median when alpha and beta are large.
- p. 5 Chapter 8 Equations 4a/4b. The *phi, lambda* parameters are parameterized on the logit and log scale respectively. The authors use the Greek letters ϕ and λ to represent these values on the logit and log scale. It is somewhat confusing to have *phi* and ϕ and *lambda* and λ used in the notation. Different covariates could be used for the two parameters, so the X value has a further subscript 1 or 2.
- p. 5 Chapter 8 Equation 5. The set of pairs of ϕ and λ come from a multi-variate normal distribution with a mean and covariance matrix.
- P.5 Chapter 8 Equation 5. The multivariate mean and variance are NOT estimated values but rather unknown population parameters that must be estimated.
- Chapter 8 Table 1 shows that the parameters of the multi-variate distribution have a posterior distribution, so a prior is needed for the multivariate parameters as well. This is not given.
- So *phi*$_y$ and *lambd*a$_y$ are derived from a hierarchical distribution. This gives a smooth cumulative distribution. We want to add "error" to this curve. This is tricky, because the cumulative distribution must add to 1. So, if in a particular week, the cumulative proportion is slightly higher than the beta curve, subsequent deviations all cannot be positive otherwise it cannot

add to 1. Chapter 8 Equation 1 purports to do this. Unfortunately, the proposed method could lead to cumulative curves that are not monotonic.

- Chapter 8 Equation 8 is not correct. The paper uses *log(p)* but *logit(p)* is likely more appropriate.
- Fitting as described in the paper is not correct as it treats each of the individual proportions as independent normal variation etc.
- The model samples from the multivariate distribution to give *phi* and *lambda,* which generate a 53-week vector of proportions that add to 1. Additional variation is also required, so a vector of size 53 that adds to 0 but ensures that none of the individual proportions falls below 0 or above 1. The current method in Chapter 8 Equation 1 could have the cumulative proportion decline which, of course, is not possible. Not clear how this will be done.

Chapter 9 Editorial and Minor Comments

Notation Issues

- A variety of notations are used that are undefined and applied inconsistently. For example, Chapter 9 Equation 1 uses *iyr* and *ir*. Presumably, detections are modelled for individual fish ("i") in each year ("yr") at receiver location ("r"). Chapter 9 Equation 2 uses *irg[i], irgT[i], itrib.* Chapter 9 Equations 3b uses *irgT*. Chapter 9 Equations 4a use Zi and ZTi etc. Presumably r*g* refers to release-group, etc. These should be defined.
- Symbols *R* and *T* distinguish between River and Tributary releases, but this is not consistently used, i.e., sometimes *R* is used and sometimes it is dropped and implied. This should be consistently used, e.g., use *surv_R100* and *surv_T100* rather than *surv_100* and *surv_T100*.
- Subscripting is not consistently used. For example, Chapter 9 Equations 4a and 4b uses S_{bSz} but this is then referred to as *S_bSz* in the text that follows - what does *b* refer to? The text following Chapter 9 Equations 4a and 4b also refers to Sz_i where the *i* index is reintroduced. There are many more instances of notational problems.

- A different notation is used for individual fish covariables (such as measures of size) and environmental covariates (Chapter 9 Equations 4a and 4b), e.g., $CovX_{i,ir}$ – What is the need for two subscripts? The text indicates that these refer to individual and reach-specific covariates, but the data does not indicate that individual environmental covariates are measured at the fish level. What are the "reach" specific covariates (reaches are between detection points)? These are not described in the data section.

- A unified notation needs to be developed. For example, all the analyses use individual fish records. These could be indexed using *i*. For each fish, there are a number of variables associated with that fish, such as the year of release (use *y* as a short cut for *y*[*i*] to indicate the year of release for fish *i*), *g* for release group (a short cut for *g*[i]), etc. All covariates, both at the fish level, at the release group level, tributary level, etc., are all recorded for the individual fish records, and so the same notation can be used.

- Some care needs to be taken in describing parameters of the models. For example, consider the modeling for detection probabilities at the 4 receiver locations described in Chapter 9 Equation 1. Just above Chapter 9 Equation 1, the text (correctly) refers to station-specific detection probabilities, but just under Chapter 9 Equation 1 the text (incorrectly) refers to detection probabilities at "reaches" (between detection locations). Similar inconsistencies occur through the document.

- A confusing terminology used in this report is the use of "survival rate" to refer both to the survival/100 kilometers (an actual rate) and the survival probability (not a rate) for a particular reach after adjusting for the reach length. Further notes about issues in the model description are noted below.

- The model descriptions use a semi-mathematical notation that is hard to follow because of inconsistent/poor notation and informal use of terms. Proper mathematical notation should be used. See for example, Chapter 9, Chapter 6 on the level of detail needed.

- There are 5-digit histories and 3-digit capture histories depending on where fish are released. But the last 2 receiver locations are common to both release locations and may therefore have the same detection probabilities. For example, in Chapter 9 Equations 2 and 3, there are common survival probabilities for reaches in common between the two release sets. Model notation for the common reach and common detection locations is somewhat awkward. Perhaps 5 digits histories should be used for BOTH River and Tributary releases. For example, histories for fish released in the river are of the form 1xxxx where x indicates detections at the downstream locations. Histories from the Tributaries would be of the form 001xx where xx indicated detections at receiver locations 3 and 4 (which are common to all release groups). This makes discussions found in Chapter 9 Tables 7 and 8 clear and more consistent because of a common receiver detection designation.

Incomplete presentation issues

- Any Bayesian analysis consists of melding a likelihood with prior information. Neither the capture-recapture nor the movement time model descriptions present the likelihood.
- In the case of the capture-recapture model, the likelihood is based on the well-known Cormack-Jolly-Seber model. The authors assume that the reader is familiar with this model and so does not provide a description. This can lead to text that appears awkward, e.g., “Estimates of survival and detection probability also depend on χ (chi in source code), the calculated probability of not being detected after the final observed detection” which requires an in depth understanding of the CJS model or requires the reader to look at the actual source code!
- In the case of the Travel Time model, the only reference to the likelihood function is found in Chapter 9 Equation 10a and 10b which model the observed travel times. The text underneath these equations indicate that a single fish could be used twice in the likelihood, e.g., a fish with history 11010 would have a likelihood term for travel time from release to Woodson, and then again from release to Sacramento. These two values from the same fish are NOT independent because the time from release to Sacramento includes the time from release to Woodson. It would be better

to split this capture history into two terms, representing the time of travel from release to Woodson, and then the time of travel from Woodson to Sacramento. [Refer to previous comments on measuring travel time when there are multiple detections in multiple receiver locations that are all pooled to the four general locations.] This is not a likelihood per se because it depends on unknown latent variables that would need to be integrated out. Rather, this is a complete-data likelihood that depends on both visible parameters and latent variables. The latent variables are "integrated" out in the MCMC process of the Bayesian model. This needs to be explained better in the text.

- No information is presented on the prior distributions for parameters of the model. For example, Chapter 9 Equation 1 models the annual probability of detection for each receiver station using a hierarchical model, i.e., the individual annual parameter values are drawn from a (hyper) distribution with a station-specific mean and standard deviation. There is no information presented on the prior distributions used for mu_r and $sigma_r$. This information is especially important when considering the final survival and detection probabilities which are confounded. Chapter 9 Equation 2 uses a mean reach-specific parameter (S_bR_{ir}) but there are no prior distributions specified for these. Chapter 9 Equations 3a/3b use a *sigma* parameter to describe random effects in Chapter 9 Equations 2a/2b but no prior distributions are identified.
- No information is presented on model fitting? For example, were the models coded in JAGS/Stan/Nimble etc. How many chains were run? How many iterations were considered burn-in. How many sample samples were saved from the posteriors etc.
- There is very limited information on model assessment, except for one mention of Rhat. A Bayesian analysis is usually assessed using traceplots, posterior predictive plots (Bayesian p-value plots), prior/posterior overlaps etc. None of this is presented and so the final fit of the Bayesian model is unknown (see overall comments below for more details).

Specific editorial issues

- The objective states that survival from the release sites to "… to Sacramento" (i.e., to the Delta entry) is of interest. Are "Sacramento" and

"Delta entry" synonymous? It may be preferable to be consistent in usage. For example, Chapter 9 Figures 9 and 15 use "Sacramento"; Chapter 9, Section 1 uses "Delta entry"; Section 2.1 uses "Delta entry". It is not until the end of page 2 that the report indicates "The JPE model only requires estimates of survival to Sacramento, which is very close to Delta entry" indicating that in fact these are two different locations but will be considered to be the same – how far apart are they?

- There are many cases where the references to Figures are incorrect. For example, in Chapter 9, Section 3.2, references to Chapter 9 Figure 7 should be to Chapter 9 Figure 10 etc. These need to be repaired.
- "Survival rate" vs "Survival". Be careful and be consistent in usage of survival rate (per 100 kilometers) and survival probability (for a reach after adjusting for length).
- p. 2 last paragraph. Reference to Chapter 9 Figure 2 should be to Figure 1.
- p. 4 "… estimated detection probabilities will be more strongly influenced by the hyper-parameters". Poorly worded. The estimates in cases of sparse data are more influenced by the prior distributions for the hyper-parameters, and not the parameters themselves.
- p. 4 last line. Reference to Chapter 9 Table 1 should point to Table 2.
- p. 4 R and T are used to distinguish between River and Tributary sections. Why not continue that notation for *surv_R100* and *surv_T100*?
- p. 4 Chapter 9 Equation 2a. *R* is used to indicate the parameters are for River released fish but *R* is not used in the *surv_100* parameter name. Be consistent. This occurs throughout the document.
- p. 4 Chapter 9 Equation 2. Survival is measured per 100 kilimeters (as a rate). The adjustment to the survival probability for a reach is given in Chapter 9 Equations 6a and 6b. There is confusion between survival rate (per 100 km) and survival probability (adjusted to actual length of each reach).
- p. 5 "Initially survival rates to Sacramento and to DeltaExit were assumed to be the same because of the low fish count at Delta exit in some years". Does this mean that the survival per 100 kilometers was assumed to be

equal in the two reaches, or the actual length adjusted survival was assumed to be equal? And be specific about the two reaches, i.e., the Sacramento to Delta reach, and the ??-Sacramento reach? “…low fish counts” should likely read “low number of fish detections”. Any low count does not imply equal survival probabilities; rather it indicates an inability to distinguish between them.

- p. 5 “Initially survival rates to Sacramento and to Delta exit were assumed to be the same because of the low fish count at Delta exit in some years”. Again, “low fish counts” should read “small number of detections”. See previous comment above.
- p. 5 “… fixed covariate effects …”. There are many types of covariates. For example, individual fish covariates that do not change over the lifespan of a fish such as sex. A pseudo-fixed individual covariate is size/weight/condition factor at the time of release. But these covariates vary over the fish’s lifetime. Then there are environmental covariates that are common for all fish in a year. And then there are environmental covariates that are common to all fish at a particular time of year, such a flow, but these certainly are not “fixed”. Suggest the report avoids use of “fixed-covariates” and instead refer to individual-covariates, year-specific covariates, and space-time covariates, etc.
- p .6 Chapter 9 Equations 4a and 4b. Change of notation between the two parts, e.g., “*Sb*” (Chapter 9 Equations 4a) and “*S_b*” (Chapter 9 Equation 4b). Also, notation differs from Chapter 9 Equations 2a and 2b. Random effect notation changes between Chapter 9 Equations 4a and 4b. Be consistent.
- p. 6 Chapter 9 Equations 5a/5b. The $S\_bCov_{1:j}$ term is presumably for a categorical variable and so needs a series of indicator variables because there are 3 levels. Notation for intercept and random effects changes between Chapter 9 Equations 5a and 5b.
- p. 6 “…were both similar across reach for a given year, and were added to the fish size effect as follows:”. What does this mean? In Chapter 9 Equations 5a and 5b, these terms are entered as $S\_bCov_{1:j}$.What is the “j” subscript – on page 7 the “*J*” is NUMBER of levels of the covariate rather than an index to the category level. Presumably this means that a set of

indicator variables are created for categorical effects that are year-specific and common to all reaches and fish – but see comments on page 7.

- p. 6 Chapter 9 Equations 4a, 4b, 5a, 5b. What are the prior distributions for the coefficients of the covariates effects?
- p. 7 "… survival rate for each reach used in the model fitting is computed using …". The conversion of the survival rate/100 km to the actual survival probability in the reach for the km travelled implies the term is no longer a "survival rate" anymore but a survival probability.
- p. 7 Chapter 9 Table 6. Too many digits presented. What is "se_mean" because the standard error is not used in Bayesian models with the SD measuring uncertainty.
- p. 7 "For j = 3, S_bCov = 0 if the individual was released in a C year, 1 if released in a D/BN year and 2 if released in a AN/W year." So, these are NOT categorical variables but are treated as pseudo-continuous with the effect of an AN/W year being twice as large as a D/BN year. Is this what is wanted? Similar comment about the flow exceedance index.
- p. 8 The final reach is denoted as "Sacramento-Delta". But Sacramento and Delta entry are very close and not distinguished in this report. Perhaps Sacramento to Delta Exit is what is meant?
- p. 10 "… Sacramento individual with a capture history of 11010, the likelihood would only be applied to observations of travel time from release to Woodson and release to Sacramento." Not sure what this means. Chapter 9 Equations 7a and 7b refer only to INDIVIDUAL reaches. So, was does travel time from release to Sacramento used in Chapter 9 Equations 7a and 7b?
- p 11 Where is the code and data files mentioned in Chapter 9, Section 2.4 R Codes?
- p. 12 "… since this was the through-Delta migration reach". The term "through-Delta" is never defined. Presumably the Sacramento to Delta Exit reach is what is meant? Be consistent in the report.
- p. 13 Reference to Chapter 9 Figure 7 should be to Figure 10.

- p. 13 “It is important to note that in a CJS model detection probability is unidentifiable from survival at the last location.” “Survival” occurs in reaches; “Detections at locations”. Reword, only the product of the survival in the final reach and detection probability at the final location is identifiable as noted earlier.
- p. 13 ”… detection rates …”. These are not rates and only a probability.
- p. 13 (and elsewhere). Be consistent in the names used for the detection locations. Here the report uses the term “Delta detection probability”; likely “Delta Exit detection probability” is better as used later on in the same paragraph.
- p. 14 “Overall, upper Sacramento River fish were faster during wetter conditions, taking as little as 3.9 days to reach Sacramento in 2017 and as much as 8.5 days in 2016 (Figure 11).” The estimates from the Bayesian models are the “mean” travel times and do not represent individual fish. Reword.
- p. 14 References to Chapter 9 Figure 11 for travel times should refer to Figure 12.
- Chapter 9 Table 1. Some release sites have data from multiple years. A chart showing for each release site which years have data would be useful.
- Chapter 9 Table 4. Please include the range in tagged fish size for each tag group. It is important to highlight the smaller fish within these tag-fish groups.
- Chapter 9 Table 5. Need to define terms used in the first column of the table, i.e. what does *MaxFlow*_FL mean? What does *WY2* mean? What does *WY3* mean?
- Chapter 9 Table 6. It is a Bayesian model. What does “*se_mean*” represent? Bayesian models report uncertainty in a parameter estimate using the SD of the posterior. How is the “confounding” in the product of the survival in the last interval and detection probability in the last receiver location accounted for?
- Chapter 9 Table 6. What are the parameters listed in column 1. Each table should stand on its own with a complete description of all notations used

in the table. For example, is “S_bSz” the coefficient for fork length? Is “S_bCov”/”S_bCovT” the coefficient for peak monthly flow?

- Chapter 9 Figure 1. Add locations of RST so that the use of the release location in Figure 1 to represent survival/travel times for fish from the RST is somewhat justified.
- Chapter 9 Figure 5 (and other figures). The 50% credible intervals are shown. Why not the usual 95% credible intervals or violin plots showing the whole posterior distribution? Yellow is a poor color choice because it is very hard to find on the white page – use a different color. The figure legend should indicate that this is from the survival model. Use better names for the reaches, e.g., Rel to Woodson; Woodson to Butte; Butte to Sacramento; and Sacramento to Delta Exit. Draw a horizontal line at 0 so we can identify release groups that consistently do better/worse than average.
- Chapter 9 Figure 5. Release groups have additional structure (year and sites and reaches). Plot these random effects by year and by sites and by reaches to see if there are any patterns. This could help determine whether models with structured random effects are worthwhile.
- Chapter 9 Figure 6. Add horizontal lines for the hierarchical mean detections. “Delta” for the location should read “DeltaExit”. Detection probabilities are usually very high except for a few cases, e.g., 2021 at Woodson bridge etc. These need further explanation.
- Chapter 9 Figure 7. Are these plots from the best fitting model as well? What values of the size and environmental covariates are used? These plots will be almost identical to those in Chapter 9 Figure 5 because they add the mean survival rate (logit scale) to the random effects seen in Chapter 9 Figure 5 and then transform to the [0,1] scale. Show the common mean on the plots (i.e. the *S_bR* from Chapter 9 Equations 2a/2b) so that we can identify release groups with better/worse survival rates compared to the average. Identify that these results are from the “null model”.
- Chapter 9 Figure 7. Different reach names used compared to Chapter 9 Figure 5. Be consistent.

- Chapter 9 Figures 7 and 8. Chapter 9 Figure 5 had both River and Tributary releases in a single plot. Now they are separated. Be consistent.
- Chapter 9 Figures 7 and 8. Release groups have additional structure (year and sites). Plot these random effects by year and sites to see if there are any patterns.
- Chapter 9 Figure 9. What model produced these estimates? The best fitting model (Chapter 9 Table 5) uses maximum flow as a continuous covariate and fork length. What value of maximum flow was used? How was this divided into wet/dry water years? What fork-length was used to make predictions (89 millimeters like in Chapter 9 Figure 10?). The figure uses "survival rate" but this is confusing because it has been adjusted for the reach lengths and should be labelled as simply survival.
- Chapter 9 Figure 9. There is additional structure (site and year). Plots by these variables could again be useful.
- Chapter 9 Figures 10 and 11 are respectively titled "*Release to Sacramento Survival Predictions, Higher Flow Values*" and "*Release to Sacramento Survival Predictions, Middle Flow Values*", but the range of flows in the description of each appears identical. Chapter 9 Figure 10 has flow on the x-axis while Figure 11 has fork length on the x-axis. Are the description and Chapter 9 Figure 11 appropriately matched?
- Chapter 9 Figure 11. Why are there no plots of estimated survival probabilities as in Chapter 9 Figure 12?
- Chapter 9 Figure 11. Modelled effects were extrapolated down to fish as small as 10 millimeter. As the MT knows, most Chinook salmon emerge from gravel at sizes of approximately 30 millimeters or greater, and none will be in the smaller sizes shown in this figure. The text identified values for 39 millimeter fish, which is also below the smallest tagged fish (70 millimeter?). Please fix this graphic oversight.
- Chapter 9 Figure 12. References to Chapter 9 Figure 10 should be to Figure 13.
- Chapter 9 Figure 12/13. Again, additional structure is available (sites, year, reaches). Some additional plots may be helpful.

- Chapter 9 Figure 14 has additional structure (site and years). Does this help in the modeling?

Appendix B - Example of Bayesian P-value

SRJPE Review Panel

1. Introduction

Posterior predictive plots (PPPs a.k.a. Bayesian p-values) are graphical tools in Bayesian statistics that check model fit by comparing a goodness-of-fit statistic based on observed data to the same goodness-of-fit statistic based simulated data generated from the posterior distribution of the parameters.

Discrepancies in the distribution of these posterior-based goodness-of-fit statistic indicate potential shortcomings in the model. The idea behind posterior predictive checking is simple: if a model is a good fit then we should be able to use it to generate data that looks a lot like the data we observed. Conversely, if a model is poor fit, then the goodness-of-fit measures computed from the observed data will be much larger than the goodness-of-fit measures computed using simulated data which, in theory, should fit the data if the model is correct. However, a good visual fit does not guarantee the model is perfect, as it may fail to capture specific, unplotted structural discrepancies.

The choice of goodness-of-fit statistic is quite general and could be an omnibus test (e.g., observed vs expected, or the deviance) or focused on more important parts of the model (what fraction of data is above a certain threshold).

In this document, we will present a simple example of a Bayesian PPP.

2. (Fake) Data generated from Efficiency Trials

We obviously don’t have the actual data used by the Modelling Team. Consequently, we generated as set of (fake) efficiency trials. This will mimic data from multiple sites over multiple years with number of releases varying in the trails, and capture probabilities around 2.5% with potential variation across weeks in a year, variation across site, and time-specific variation.

The design is balanced across sites, weeks, and year, but that plays no part in the subsequent analysis.

In the dataset below, we have a (random) Site effect, but no other random effects.

```
# Generate efficiency data.
# For this exampe, we generate data with a random site effect only.

N.Years <- 10
N.Sites <- 5
N.Weeks <- 10

year.effect.sd <- 0
site.effect.sd <- 2
site.year.effect.sd <- 0
site.year.week.effect.sd <- 0
```

Appendix B

```
eff.data <- expand.grid(Year=1:N.Years, Site=1:N.Sites, Week=1:N.Weeks)
eff.data$logit.P.base <- log(.025/(1-.025))

year.effects <- data.frame(Year=1:N.Years,
                           year.effect=rnorm(N.Years,
                                             mean=0, sd=year.effect.sd))

site.year.effects <- expand.grid(Year=1:N.Years, Site=1:N.Sites)
site.year.effects$site.year.effect = rnorm(N.Years*N.Sites,
                                           mean=0, sd=site.year.effect.sd)

site.effects <- data.frame(Site=1:N.Sites,
                           site.effect=rnorm(N.Sites,
                                             mean=0, sd=site.effect.sd))

eff.data <- merge(eff.data, year.effects)
eff.data <- merge(eff.data, site.year.effects)
eff.data <- merge(eff.data, site.effects)

eff.data$logit.P <- eff.data$logit.P.base +
  eff.data$year.effect +
  eff.data$site.effect +
  eff.data$site.year.effect +
  rnorm(nrow(eff.data), mean=0,
        sd=site.year.week.effect.sd) # just random noise at week level
eff.data$P <- 1/(1+exp(-eff.data$logit.P))

# number of fish released
eff.data$n <- rep(c(seq(20,100,40), seq(150,1000,50)),
                  length.out=nrow(eff.data))
#eff.data$n <- 500
eff.data$x <- rbinom(nrow(eff.data), size=eff.data$n, prob=eff.data$P)
```

We compute the mean observed efficiency over the weeks within a site.year as shown in Figure 4.

```
# Compute the mean observed efficiency ove weeks within a site.year
eff.data.mean.phat <- plyr::ddply(eff.data, c("Site","Year"), plyr::summarize
,
                                  mean.p.hat = mean(x/n),
                                  mean.p.hat2=sum(x)/sum(n))

ggplot(data=eff.data.mean.phat, aes(x=Year, y=logit(mean.p.hat2),
                                    color=as.factor(Site),
                                    linetype=as.factor(Site),
                                    shape=as.factor(Site)))+
   geom_point()+
   geom_line()+
   scale_color_discrete(name="Site")+
```

Appendix B

```
    scale_linetype_discrete(name="Site")+
    scale_shape_discrete(name="Site")
```

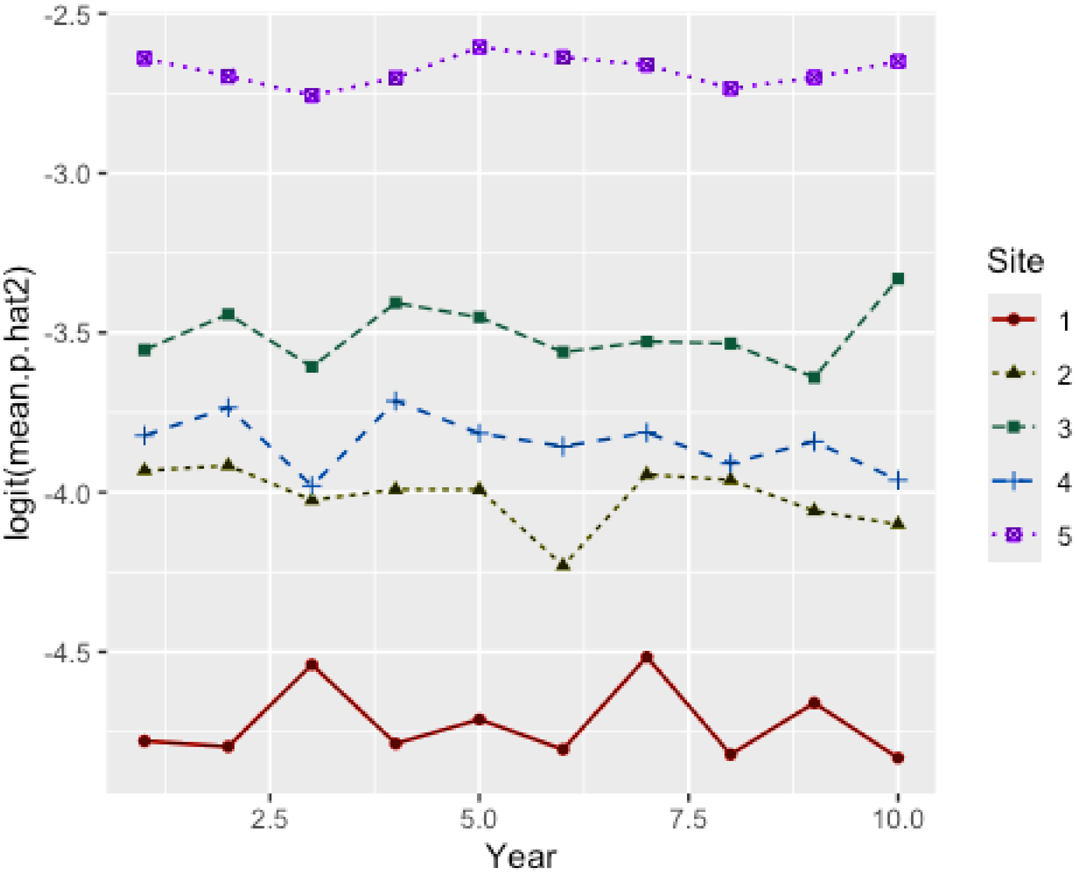


Figure 4. Mean efficiency by week.

There appears to be some structure in the data with some sites tending to have lower/higher efficiency values consistently over time.

3. Model 0 (Null model)

We fit a simple model where p is assumed to be equal across all sites, years, and weeks to illustrate the PPP when the model should be a poor fit to the data.

```
# fit the null model. Wrap it into a bigger list incase you
# want to fit multiple models (more advanced)
result0 <- list(Model=fit.model.0(eff.data))

Compiling model graph
   Resolving undeclared variables
   Allocating nodes
Graph information:
   Observed stochastic nodes: 500
   Unobserved stochastic nodes: 2
```

```
   Total graph size: 1009

Initializing model
```

3.1 Prior distributions

3.1.1 mu.logit

The prior distribution in Model 0 for mu.logit is found in Figure 5

```
# Extract the prior distribution generated by the model

mu.logit.prior <-  extract.post(result0, "mu.logit.prior")

ggplot(data=mu.logit.prior, aes(x=value, color=parameter))+
  ggtitle("Prior  for mu.logit")+
  geom_density()+xlab("mu.logot")+
  facet_wrap(~Surveys,ncol=2)
```

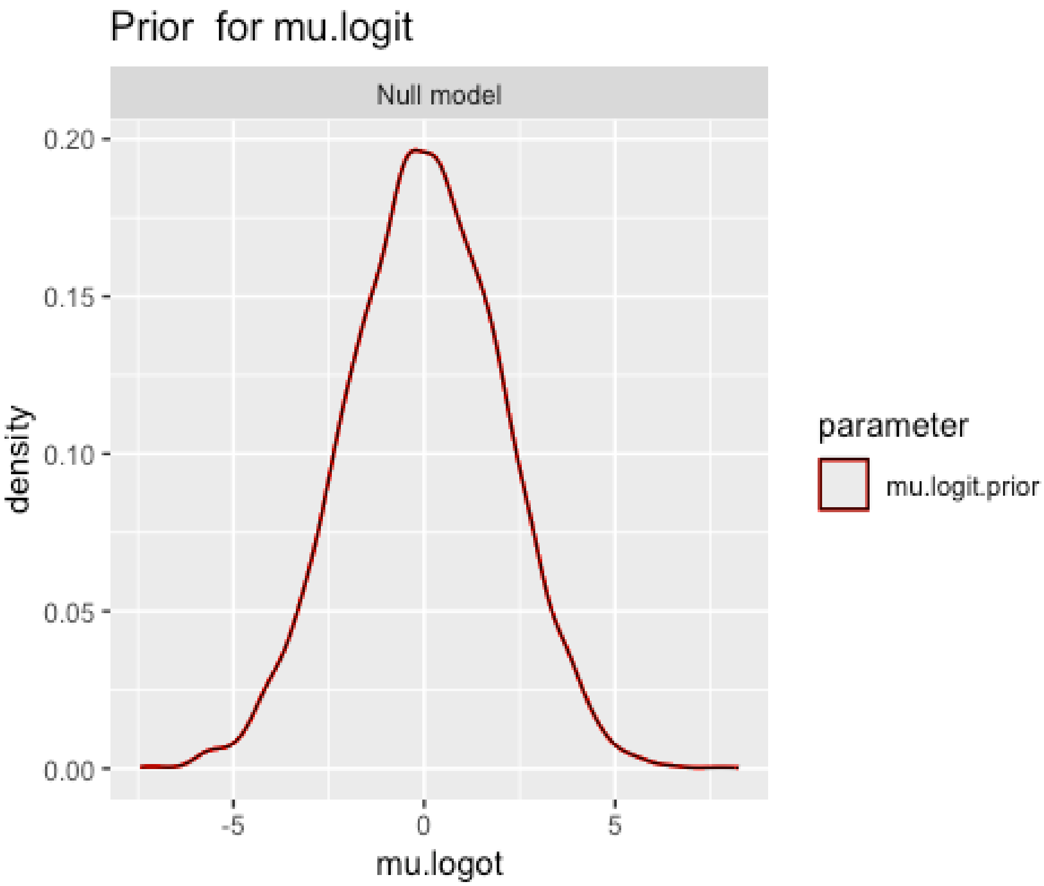


Figure 5. Prior distribution for mu.logit from Model 0.

Appendix B

3.2 Posterior: mu.logit

The posterior distribution in Model 0 for mu.logit is found in Figure 6

```
# Extract the posterior for mu.logit
mu.logit.post <-  extract.post(result0, "mu.logit$")

ggplot(data=mu.logit.post, aes(x=value, color=parameter))+
  ggtitle("Posterior  for mu.logit")+
  geom_density()+xlab("mu.logot")+
  facet_wrap(~Surveys,ncol=2)
```

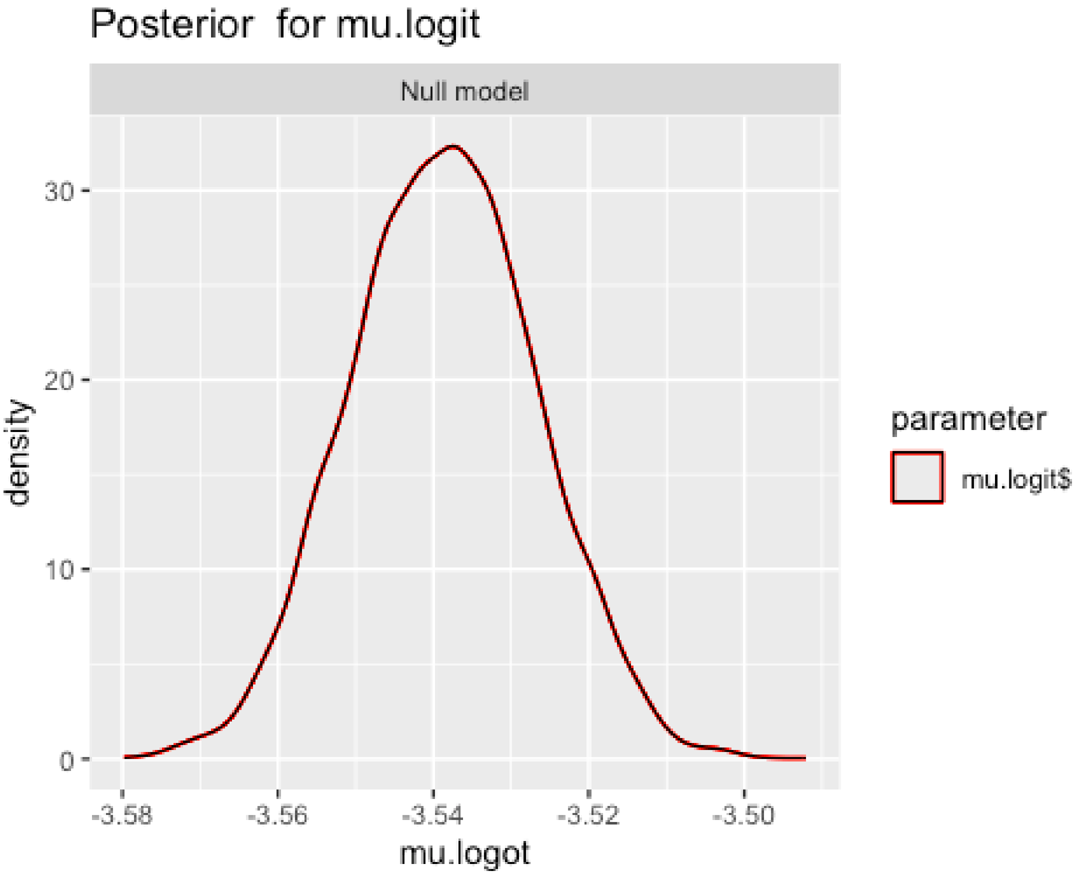


Figure 6. Posterior distribution for mu.logit from Model 0.

The prior and posterior overlap is shown in Figure 7

```
# Find the prior and posterior overlap

overlap <- data.frame(overlap=postPriorOverlap(mu.logit.post$value,
                                               mu.logit.prior$value))


temp <- plyr::rbind.fill(mu.logit.prior, mu.logit.post)
```

```
ggplot(data=temp, aes(x=value, color=parameter))+
  ggtitle("Prior and posterior for mu logit")+
  geom_density(aes(linetype=parameter))+xlab("mu logit")+
  geom_text(data=overlap, aes(label=paste0("Overlap: ",round(overlap,3)),
                                x=-Inf, y=Inf, color=NULL, hjust=0, vjust=1.5))
```

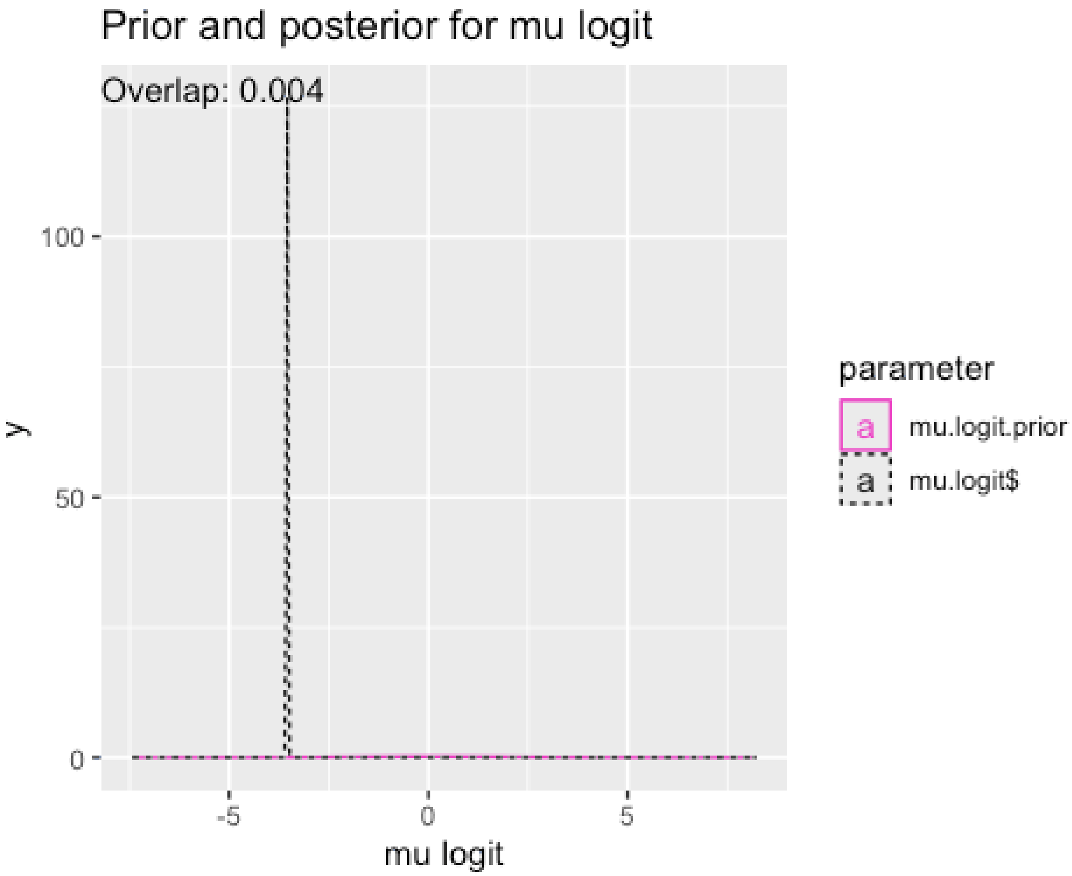


Figure 7. Posterior distribution for mu.logit from Model 0.

3.3 Posterior: p.cap

We show the summary of the first few efficiency trials – they are all identical because of the very (unrealistic) simple model.

```
# we extract the summary of the posterior distribution for
# the individual efficiency data
# we only show the first few rows.

p.cap.post.summary <-  extract.est(result0, "p.cap\\[")

old.digits <- getOption("digits")
options(digits=3)
head(p.cap.post.summary)
```

Appendix B

```
    .id    Surveys   mean       sd  X2.5. X97.5. Rhat n.eff  pattern
1 Model Null model 0.0282 0.000327 0.0276 0.0289    1  6000 p.cap\\[
2 Model Null model 0.0282 0.000327 0.0276 0.0289    1  6000 p.cap\\[
3 Model Null model 0.0282 0.000327 0.0276 0.0289    1  6000 p.cap\\[
4 Model Null model 0.0282 0.000327 0.0276 0.0289    1  6000 p.cap\\[
5 Model Null model 0.0282 0.000327 0.0276 0.0289    1  6000 p.cap\\[
6 Model Null model 0.0282 0.000327 0.0276 0.0289    1  6000 p.cap\\[

options(digits=old.digits)
```

3.4 PPP

The predictive posterior plot is found by finding a goodness-of-fit measure for the actual data and for simulated data based on the posterior values.

We first find an array of posterior samples for each of the p.cap values and 1 row for each posterior sample, pooled over all chains

```
# array of posterior values
p.cap.post.array <-  result0$Model$results$BUGSoutput$sims.list$p.cap

dim(p.cap.post.array)

[1] 6000  500
```

Next we compute a measure of goodness of fit based on observed data for each row in the above array. We will use Freeman-Tukey (Freeman and Tukey 1950) statistic that compares observed and expected counts:

$$D(x;\theta) = \sum\left(\sqrt{obs} - \sqrt{exp}\right)^2$$

As indicated by Brooks et al. (2000), this discrepancy measure is less sensitive to very small expected counts and removes the need for pooling cells often found in the classical Pearson chi-square goodness-of-fit.

```
# Compute the discrepancy measure (Freeman-Tukey) for the observed data
discrep.obs <- plyr::aaply(p.cap.post.array,1,function(x){
   # compute expected recaptures
   e.rec <- x * eff.data$n
   discrep <- sum(sqrt(e.rec) - sqrt(eff.data$x))**2
   discrep
})
```

Next we compute the same goodness of fit measure based on SIMULATED data for each row in the above array.

```
# Compute the discrepancy measure (Freeman-Tukey) for the simulated data
# using current model
discrep.sim <- plyr::aaply(p.cap.post.array,1,function(x){
   # compute expected recaptures
```

```
   e.rec <- x * eff.data$n
   # compute simulated data
   sim.x <- rbinom(length(x), size=eff.data$n, prob=x)
   discrep <- sum(sqrt(e.rec) - sqrt(sim.x))**2
   discrep
})
```

We now have two vectors of discrepancy measured. There are number of ways to compare them. See the documentation of the **`bayesplot`** package for more details.

One of the most common is the Bayesian p-value plot where the observed and simulated discrepancy measures are plotted against each other, and the proportion of time one exceeds the other is computed (Figure 8)

```
# Produce the Bayesian p-value plot

temp <- data.frame(obs=discrep.obs, sim=discrep.sim)

p.value <- mean(temp$obs > temp$sim)

ggplot(data=temp, aes(x=obs, y=sim))+
   geom_point()+
   geom_abline(intercept=0, slope=1)+
   annotate("text", label=paste0("Bayesian p-value: ", round(p.value,2)),
            x=-Inf, y=Inf, hjust=-1.5, vjust=1.5)
```

Appendix B

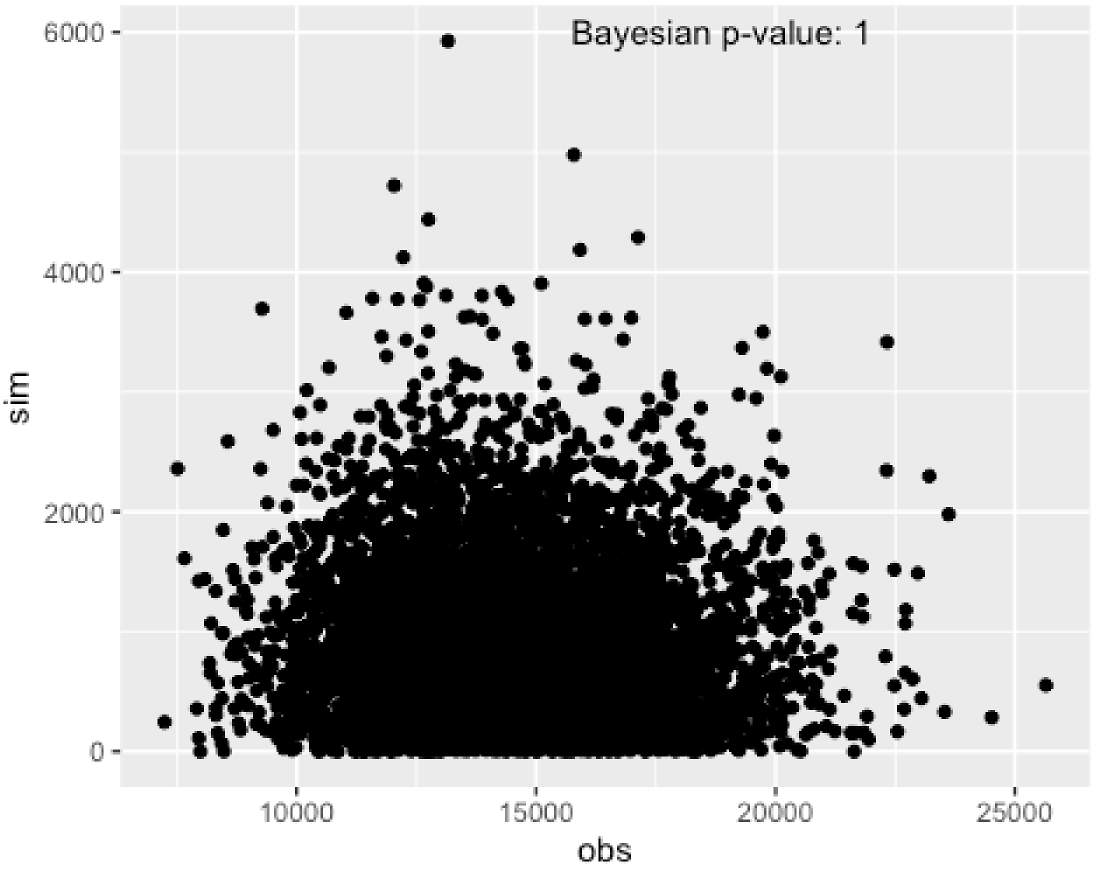


Figure 8. Bayesian posterior predictive plot (Bayesian p-value plot) for Model 0.

A Bayesian p-value that is FAR from 0.5 indicates a lack of fit.

In this case, there is evident lack of fit. Because the goodness-of-fit statistic based on the observed data is much larger than based on the simulated data, this indicates that there is more variation in the observed data than the model “explains”.

This is in accordance with the simulated data that had an additional (random) site effect.

Notice that even for this badly fitting model, values of `Rhat` are acceptable. This is the primary reason why reliance on `Rhat` alone is not sufficient for model assessment.

4. Model 1 (Random site effects)

We fit a slightly more complex model where there is (random) site effect but no consistent (random) time effects or (random) site.time effects. This model now matches the model used to generate the simulated data, so we expect a good fit.

```
# Fit the model with a random site effect

result1 <- list(Model=fit.model.1(eff.data))
```

Appendix B

```
Compiling model graph
   Resolving undeclared variables
   Allocating nodes
Graph information:
   Observed stochastic nodes: 500
   Unobserved stochastic nodes: 8
   Total graph size: 1530

Initializing model
```

4.1 Prior distributions

4.1.1 mu.logit

The prior distribution for mu.logit from Model 1 is shown in Figure 9.

```
mu.logit.prior <-  extract.post(result1, "mu.logit.prior")

ggplot(data=mu.logit.prior, aes(x=value, color=parameter))+
  ggtitle("Prior  for mu.logit")+
  geom_density()+xlab("mu.logot")+
  facet_wrap(~Surveys,ncol=2)
```

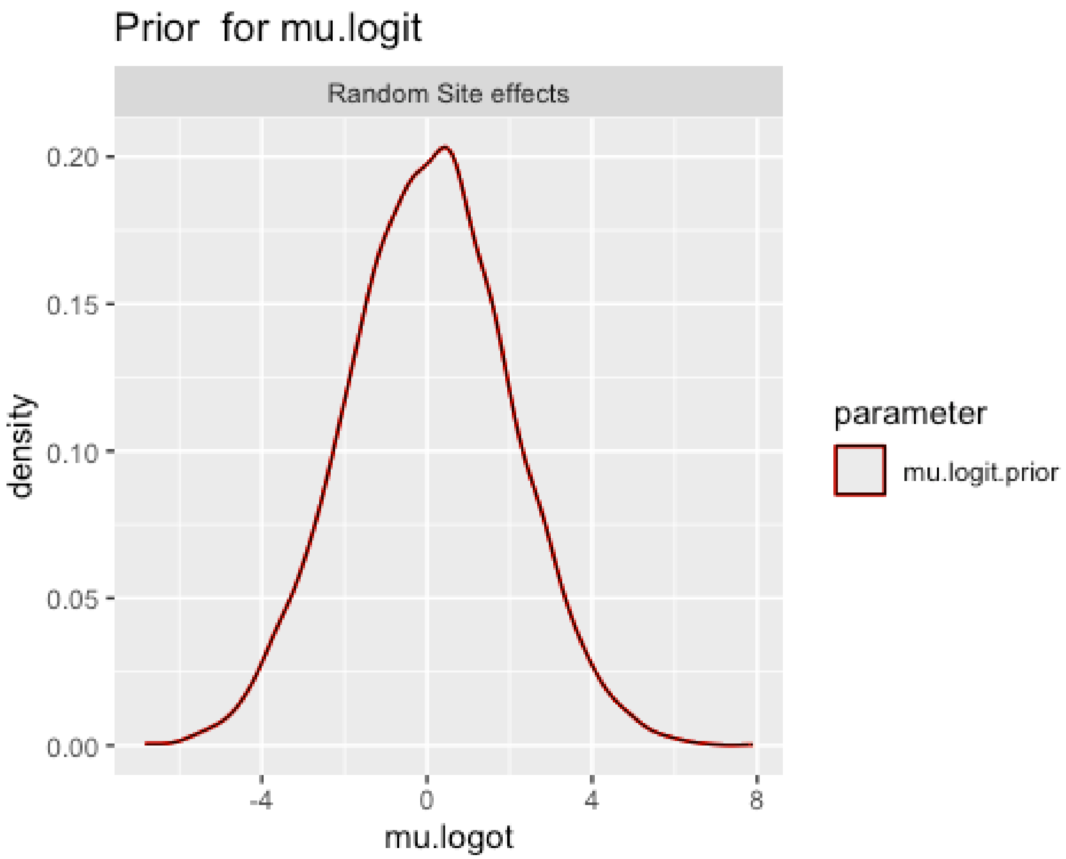


Figure 9. Prior distribution for mu.logit from Model 1.

4.2 Posterior: mu.logit

The posterior distribution in Model 1 for mu.logit is found in Figure 10

```
# Extract the posterior for mu.logit
mu.logit.post <-  extract.post(result1, "mu.logit$")

ggplot(data=mu.logit.post, aes(x=value, color=parameter))+
  ggtitle("Posterior  for mu.logit")+
  geom_density()+xlab("mu.logot")+
  facet_wrap(~Surveys,ncol=2)
```

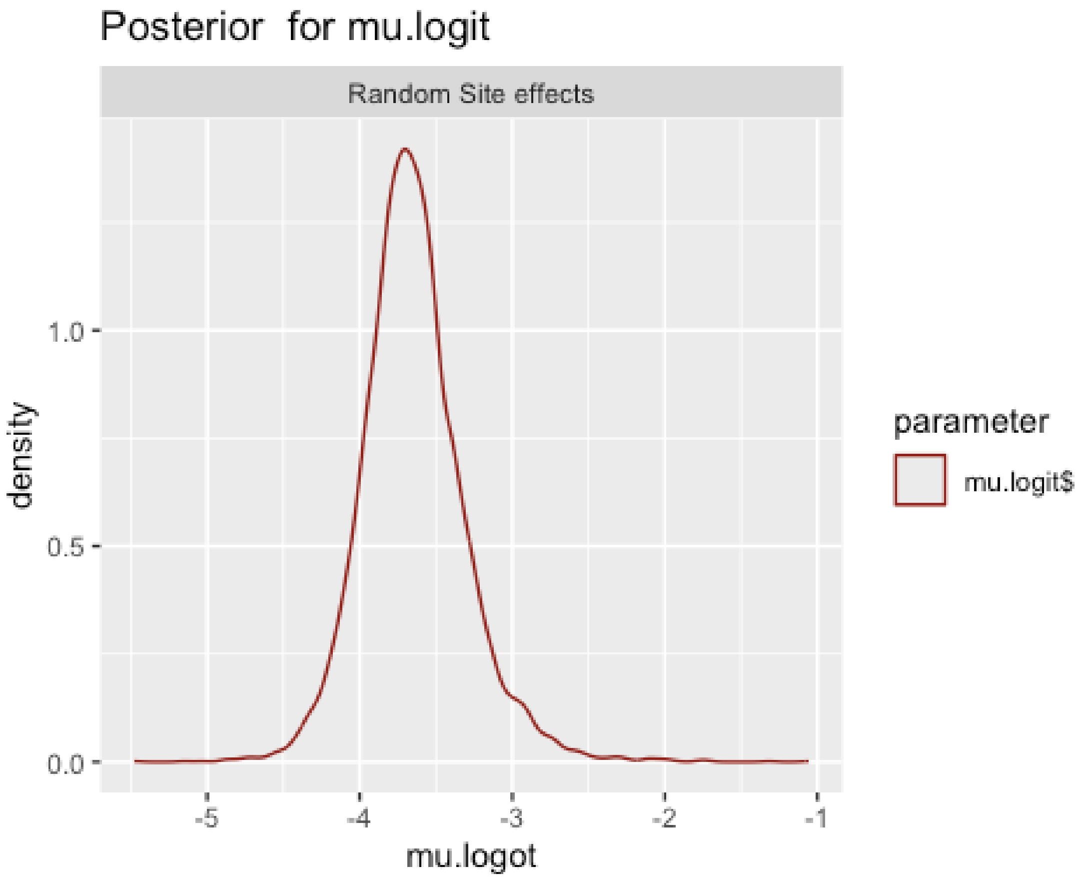


Figure 10. Posterior distribution for mu.logit from Model 1.

The prior and posterior overlap is shown in Figure 11

```
# Find the prior and posterior overlap

overlap <- data.frame(overlap=postPriorOverlap(mu.logit.post$value,
                                               mu.logit.prior$value))


temp <- plyr::rbind.fill(mu.logit.prior, mu.logit.post)

ggplot(data=temp, aes(x=value, color=parameter))+
  ggtitle("Prior and posterior for mu logit")+
  geom_density(aes(linetype=parameter))+xlab("mu logit")+
  geom_text(data=overlap, aes(label=paste0("Overlap: ",round(overlap,3)),
                              x=-Inf, y=Inf, color=NULL, hjust=0, vjust=1.5))
```

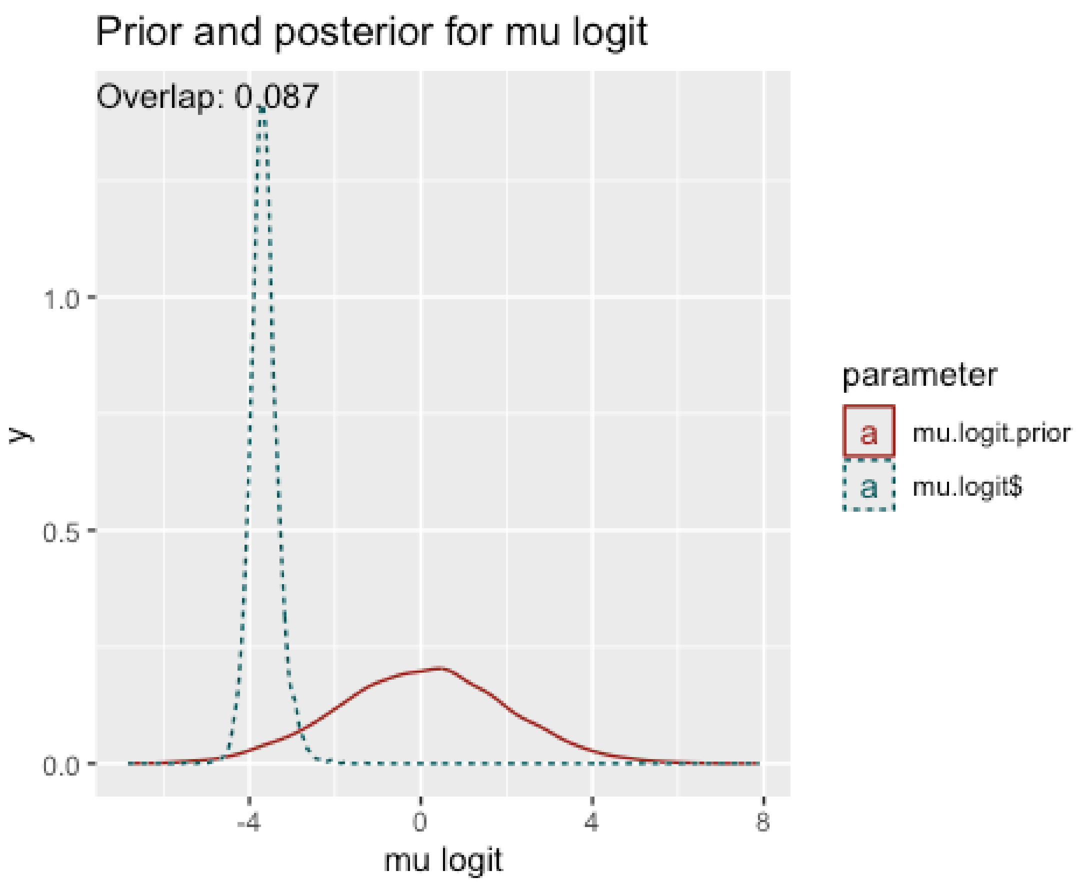


Figure 11. Prior and posterior overlap for mu.logit from Model 1.

4.3 Posterior: site effects

We show the summary of the site effects

```
# we extract the summary of the posterior distribution for the site effects

site.eff.post.summary <-  extract.est(result1, "^site.eff.logit")

old.digits <- getOption("digits")
options(digits=3)


site.eff.post.summary

    .id             Surveys   mean     sd  X2.5. X97.5. Rhat n.eff         pat
tern
1 Model Random Site effects -1.050 0.342 -1.806 -0.437    1  6000 ^site.eff.l
ogit
2 Model Random Site effects -0.359 0.340 -1.109  0.241    1  3100 ^site.eff.l
ogit
3 Model Random Site effects  0.144 0.340 -0.600  0.762    1  3300 ^site.eff.l
ogit
```

Appendix B

```
4 Model Random Site effects -0.199 0.340 -0.946  0.416    1  4800 ^site.eff.l
ogit
5 Model Random Site effects  0.970 0.339  0.222  1.580    1  3500 ^site.eff.l
ogit

""

[1] ""

sd.site.eff.post.summary <-  extract.est(result1, "^sd.site.eff.logit")
sd.site.eff.post.summary

    .id              Surveys  mean    sd X2.5. X97.5. Rhat n.eff            pa
ttern
1 Model Random Site effects 0.702 0.255 0.399   1.35    1  6000 ^sd.site.eff.
logit

options(digits=old.digits)
```

4.4 Posterior: p.cap

We show the summary of the first few efficiency trials – they are all identical because of the very (unrealistic) simple model.

```
# we extract the summary of the posterior distribution for the
# individual efficiency data
# we only show the first few rows.

p.cap.post.summary <-  extract.est(result1, "p.cap\\[")

old.digits <- getOption("digits")
options(digits=3)

head(p.cap.post.summary,n=15)

     .id              Surveys    mean       sd   X2.5.   X97.5. Rhat n.eff  pat
tern
1  Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
2  Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
3  Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
4  Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
5  Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
6  Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
7  Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
```

```
8  Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
9  Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
10 Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
11 Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
12 Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
13 Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
14 Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
15 Model Random Site effects 0.00904 0.000423 0.00824 0.00991 1.01   220 p.ca
p\\[
```

```
""
```

```
[1] ""
```

```
tail(p.cap.post.summary,n=15)
```

```
      .id             Surveys   mean      sd  X2.5. X97.5. Rhat n.eff  patter
n
486 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
487 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
488 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
489 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
490 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
491 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
492 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
493 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
494 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
495 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
496 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
497 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
498 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
```

```
499 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[
500 Model Random Site effects 0.0643 0.00109 0.0622 0.0665    1  1100 p.cap\\
[

options(digits=old.digits)
```

Notice that the p.cap values are grouped by site (10 values for each site) and now vary by site, but not over year in the simulated data.

4.5 PPP

The predictive posterior plot is found by finding a goodness-of-fit measure for the actual data and for simulated data based on the posterior values.

We first find an array of posterior samples for each of the p.cap values and 1 row for each posterior sample, pooled over all chains

```
# array of posterior values
p.cap.post.array <-  result1$Model$results$BUGSoutput$sims.list$p.cap

dim(p.cap.post.array)

[1] 6000  500
```

Next, we compute a measure of goodness of fit based on observed data for each row in the above array. We will use Freeman-Tukey (Freeman and Tukey 1950) statistic that compares observed and expected counts:

$$D(x;\theta) = \sum\left(\sqrt{obs} - \sqrt{exp}\right)^2$$

As indicated by Brooks et al. (2000), this discrepancy measure is less sensitive to very small expected counts and removes the need for pooling cells often found in the classical Pearson chi-square goodness-of-fit.

```
# discrepancy value using observed data (Freeman-Tukey statistics)
discrep.obs <- plyr::aaply(p.cap.post.array,1,function(x){
   # compute expected recaptures
   e.rec <- x * eff.data$n
   discrep <- sum(sqrt(e.rec) - sqrt(eff.data$x))**2
   discrep
})
```

Next, we compute the same goodness of fit measure based on SIMULATED data for each row in the above array.

```
# discrepancy value using simulated data (Freeman-Tukey statistics)
discrep.sim <- plyr::aaply(p.cap.post.array,1,function(x){
   # compute expected recaptures
   e.rec <- x * eff.data$n
   # compute simulated data
```

```
   sim.x <- rbinom(length(x), size=eff.data$n, prob=x)
   discrep <- sum(sqrt(e.rec) - sqrt(sim.x))**2
   discrep
})
```

We now have two vectors of discrepancy measured. There are number of ways to compare them. See the documentation of the `bayesplot` package for more details.

One of the most common is the Bayesian p-value plot where the observed and simulated discrepancy measures are plotted against each other, and the proportion of time one exceeds the other is computed (Figure 12)

```
# Bayesian p-value plot
temp <- data.frame(obs=discrep.obs, sim=discrep.sim)

p.value <- mean(temp$obs > temp$sim)

ggplot(data=temp, aes(x=obs, y=sim))+
   geom_point()+
   geom_abline(intercept=0, slope=1)+
   annotate("text", label=paste0("Bayesian p-value: ", round(p.value,2)),
            x=-Inf, y=Inf, hjust=-1.5, vjust=1.5)
```

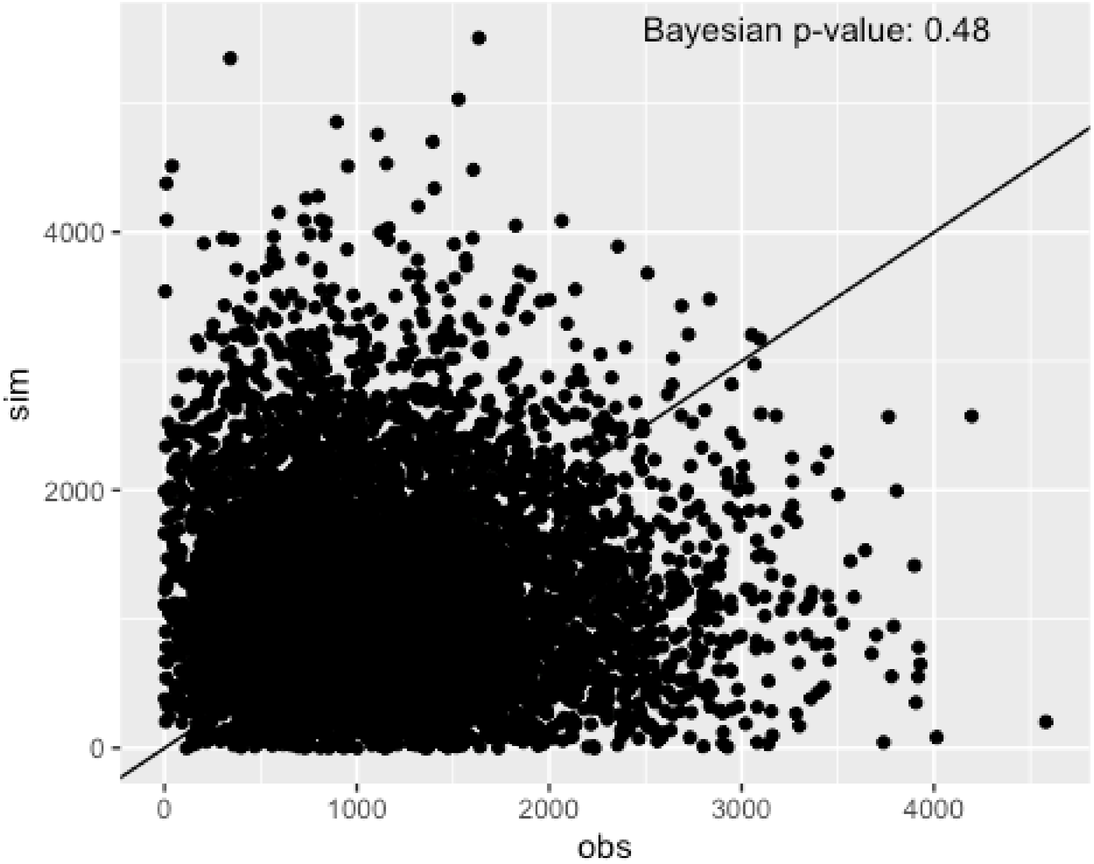


Figure 12. Bayesian posterior predictive plot (Bayesian p=value plot) for Model 1.

A Bayesian p-value that is FAR from 0.5 indicates a lack of fit.

In this case, the fit is improved compared to the null model (as expected).

5. References

Brooks, S.P., E.A. Catchpole and B.J.T. Morgan. (2000). Bayesian animal survival estimation. *Statistical Science* 15(4), 357-376. http://www.jstor.org/stable/2676830

Freeman, M. F. and J.W. Tukey. (1950). Transformations related to the angular and square root. *Annals of Mathematical Statistics* 21(4), 607-611. https://doi.org/10.1214/aoms/1177729756

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). Bayesian Data Analysis. Chapman & Hall/CRC Press, London, third edition. (Ch. 6). https://doi.org/10.1201/b16018

Gimenez, O., Morgan, B.J., and Brooks, S. (2009). Weak identifiability in models for mark-recapture-recovery data. pp.1055-1068 in Thomson, Cooch and Conroy (eds) *Modeling demographic processes in marked populations*, vol 3. Springer, Boston, MA.. https://doi.org/10.1007/978-0-387-78151-8_48

Lunn, D., Jackson, C., Best, N., Thomas, A. and Spiegelhalter, D. (2012). *The BUGS Book – A practical introduction to Bayesian Analysis*. Chapman and Hall/CRC Press. https://doi.org/10.1201/b13613

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. http://citeseer.ist.psu.edu/plummer03jags.html

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

6. Function listings

```
# Fit hierarchical Bayesian models to the efficiency data

###############################################################################
##############################
###############################################################################
#############################

# Model 0 is the null model with a common mean over all sites and years
# and no random effects.
fit.model.0 <- function(eff.data){

# Variables in model are
#  N.eff.trials - number of efficiency trial
#  n.rel        - number of fish released
#  x.rec        = number of fish released that are recovered

# The BUGS model is specified as a text file.

# The model file.
# The cat() command is used to save the model to the working directory.
# Notice that you CANNOT have any " (double quotes) in the bugs code
# between the start and end of the cat("...",) command.

cat(file="model.txt", "
############################################################

model {
  #  N.eff.trials - number of efficiency trial
  #  n.rel    - number of fish released vector of length N.eff.trials
```

```
  #  x.rec   - number of fish released that are recovered of length N.eff.tri
als

  # number of recaptures modelled as binomial distribution
  for(i in 1:N.eff.trials){
     x.rec[i] ~ dbin( p.cap[i], n.rel[i])
  }

  # convert logit to p
  for(i in 1:N.eff.trials){
     p.cap[i] <- ilogit(p.cap.logit[i])
  }

  # model on logit scale
  for(i in 1:N.eff.trials){
     p.cap.logit[i] <- mu.logit
  }

  # prior on mu.logit (roughly flat)
  mu.logit       ~  dnorm(0, 1/(2^2))  # rougly uniform between 0 and 1
  mu.logit.prior ~  dnorm(0, 1/(2^2))

  mu.p     <- ilogit(mu.logit)

  #tau.logit ~ dunif(.01, 2)
  #sd.logit <- 1/sqrt(tau.logit)
}
") # End of the model

  N.eff.trials <- nrow(eff.data)

  #  Next create the data.txt file.
  # Initialize the data values using standard R code by either reading
  # in from an external file, or plain assignment.

  # The datalist will be passed to JAGS with the names of the data
  # values.
  data.list <- list(N.eff.trials        = nrow(eff.data),
                    n.rel               = eff.data$n,
                    x.rec               = eff.data$x
                  ) # initial prior for sex/age breakdown

  # check the list
  data.list


  # Next create the initial values.

  init.list <- list(
      list(mu.logit=log(.025/(1-.025))),
```

```
      list(mu.logit=log(.025/(1-.025))),
      list(mu.logit=log(.025/(1-.025)))
      )  # end of list of lists of initial values

  # Next create the list of parameters to monitor.
  #
  monitor.list <- c("mu.logit", "p.cap",
                    "mu.logit.prior")


  results <- jags(
      data      =data.list,   # list of data variables
      inits     =init.list,   # list/function for initial values
      parameters=monitor.list,# list of parameters to monitor
      model.file="model.txt",  # file with bugs model
      n.chains=3,
      n.iter  =22000,          # total iterations INCLUDING burn in
      n.burnin=2000,           # number of burning iterations
      n.thin=10,               # how much to thin
      DIC=FALSE,               # is DIC to be computed?
      working.dir=getwd()    # store results in current working directory
      )

  list(name="Null model",
       results=results)
}

#############################################################################
##############################
#############################################################################
#############################
# Model 1 includes a random site effect over and above the null model.

fit.model.1 <- function(eff.data){

  # Variables in model are
  #  N.eff.trials - number of efficiency trial
  #  N.Sites      - number of sites for the random site effects
  #                 (on the logit scale)
  #                 Sites must be numbered 1...N.Sites
  #  n.rel        - number of fish released
  #  x.rec        = number of fish released that are recovered

  # The BUGS model is specified as a text file.

  # The model file.
  # The cat() command is used to save the model to the working directory.
  # Notice that you CANNOT have any " (double quotes) in the bugs code
  # between the start and end of the cat("...",) command.
```

```
  cat(file="model.txt", "
############################################################

model {
  #  N.Sites      - number of sites (numbered 1..N.Sites)
  #  N.eff.trials - number of efficiency trials
  #  n.rel        - number of fish released vector of length N.eff.trials
  #  x.rec        - number of fish released that are recovered of
  #                 length N.eff.trials
  #  Site         - site number of each trial

  # number of recaptures modelled as binomial distribution
  for(i in 1:N.eff.trials){
     x.rec[i] ~ dbin( p.cap[i], n.rel[i])
  }

  # convert logit to p
  for(i in 1:N.eff.trials){
     p.cap[i] <- ilogit(p.cap.logit[i])
  }

  # Site effects
  for(i in 1:N.Sites){
     site.eff.logit[i] ~ dnorm(0, tau.site.eff.logit)
  }

  tau.site.eff.logit ~dunif(.01,8)
  sd.site.eff.logit <- 1/sqrt(tau.site.eff.logit)


  # model on logit scale
  for(i in 1:N.eff.trials){
     p.cap.logit[i] <- mu.logit + site.eff.logit[Site[i]]
  }

  # prior on mu.logit (roughly flat)
  mu.logit       ~  dnorm(0, 1/(2^2))  # rougly uniform between 0 and 1
  mu.logit.prior ~  dnorm(0, 1/(2^2))

  mu.p     <- ilogit(mu.logit)

}
") # End of the model

  N.eff.trials <- nrow(eff.data)
  N.Sites      <- max(eff.data$Site)
  #browser()

  #  Next create the data.txt file.
  # Initialize the data values using standard R code by either reading
```

```
  # in from an external file, or plain assignment.

  # The datalist will be passed to JAGS with the names of the data
  # values.
  data.list <- list(N.eff.trials        = nrow(eff.data),
                    n.rel               = eff.data$n,
                    x.rec               = eff.data$x,
                    N.Sites             = N.Sites,
                    Site                = eff.data$Site
  ) # initial prior for sex/age breakdown

  # check the list
  data.list


  # Next create the initial values.

  init.list <- list(
    list(mu.logit=log(.025/(1-.025))),
    list(mu.logit=log(.025/(1-.025))),
    list(mu.logit=log(.025/(1-.025)))
  )  # end of list of lists of initial values

  # Next create the list of parameters to monitor.
  #
  monitor.list <- c("mu.logit", "p.cap", "sd.site.eff.logit","site.eff.logit"
,
                    "mu.logit.prior")


  results <- jags(
    data      =data.list,   # list of data variables
    inits     =init.list,   # list/function for initial values
    parameters=monitor.list,# list of parameters to monitor
    model.file="model.txt",  # file with bugs model
    n.chains=3,
    n.iter  =22000,          # total iterations INCLUDING burn in
    n.burnin=2000,           # number of burning iterations
    n.thin=10,               # how much to thin
    DIC=FALSE,               # is DIC to be computed?
    working.dir=getwd()    # store results in current working directory
  )

  list(name="Random Site effects",
       results=results)
}



#############################################################################
```

```
##############################
#############################################################################
#############################


# extract estimates and 95% ci from the fitted objects
extract.est <- function (all.fits, pattern, n.dec=2){
  # pattern identifies the estimates to extract using grepl
  res <- plyr::ldply(all.fits, function(x){
    #browser()
    temp <- x$results$BUGSoutput$summary[grepl(pattern,
                                      row.names(x$results$BUGSoutput$summary)
)
    ,c("mean", "sd", "2.5%","97.5%","Rhat", "n.eff"), drop=FALSE]
    names(temp)<- c("Estimate","SD", "lci.95", "uci.95", "Rhat","n.eff")
    temp <- data.frame(temp)
    temp <- cbind(Surveys=x$name, temp)
    temp$pattern <- pattern
    temp$Rhat <- round(temp$Rhat,5)
    temp$n.eff<- round(temp$n.eff,0)
    # get the 90% ci
    #browser()
    #temp2 <- HDInterval::hdi(x$results, credMass=0.90)
    #temp2 <- temp2[,grepl(pattern,colnames(temp2 ))]
    #temp2 <- t(temp2)
    #temp$lci.90 <- round(temp2[,"lower"],n.dec)
    #temp$uci.90 <- round(temp2[,"upper"],n.dec)
    temp
  })
  res
}

#extract.est(all.fits, "p.alive$")
#extract.est(all.fits, "p.detect")  # multiple values

#############################################################################
##############################
#############################################################################
#############################
# extract posterior samples
extract.post <- function (all.fits, pattern){
  # pattern identifies the estimates to extract using grepl
  res <- plyr::ldply(all.fits, function(x){
    #browser()
    select <- names(x$results$BUGSoutput$sims.list)[
                    grepl(pattern,names(x$results$BUGSoutput$sims.list))]
    if(length(select)>1)stop("extract.post --pattern selects more than one ",
                             "parameter. pattern:",pattern,
                             "; selected ", knitr::combine_words(select))
    temp1 <- pivot_longer(as.data.frame(x$results$BUGSoutput$sims.list[[selec
```

Appendix B

```
t]]),
                             cols=everything())
    temp1$Surveys=x$name
    temp1$parameter=pattern
    temp1
  })
  res
}

#temp <- extract.post(all.fits, "^p.alive$")
#temp <- extract.post(all.fits, "^p.detect")



##############################################################################
##############################
##############################################################################
#############################
# estimate the overlap between the prior and posterior distribution

postPriorOverlap <- function (paramSampleVec, prior, ..., yaxt = "n", ylab =
"",
    xlab = "Parameter", main = "", cex.lab = 1.5, cex = 1.4,
    xlim = range(paramSampleVec), breaks = NULL, plot=FALSE)
{
  # This is a copy of the postPriorOverlap function from the BEST package
  # except that we stopped the display of the histograms on the plot window
  require(HDInterval)
    oldpar <- par(xpd = NA)
    on.exit(par(oldpar))
    if (is.null(breaks)) {
        nbreaks <- ceiling(diff(range(paramSampleVec))/
                             as.numeric(diff(hdi(paramSampleVec))/18))
        breaks <- seq(from = min(paramSampleVec), to = max(paramSampleVec),
            length.out = nbreaks)
    }
    histinfo <- hist(paramSampleVec, xlab = xlab, yaxt = yaxt,
        ylab = ylab, freq = FALSE, border = "white", col = "skyblue",
        xlim = xlim, main = main, cex = cex, cex.lab = cex.lab,
        breaks = breaks, plot=plot)
    if (is.numeric(prior)) {
        priorInfo <- hist(prior, breaks = c(-Inf, breaks, Inf),
            add = TRUE, freq = FALSE, col = "yellow", border = "white",
            plot=plot)$density[2:length(breaks)]
    }
    else if (is.function(prior)) {
        if (class(try(prior(0.5, ...), TRUE)) == "try-error")
            stop(paste("Incorrect arguments for the density function",
                substitute(prior)))
        priorInfo <- prior(histinfo$mids, ...)
```

```
    }
    minHt <- pmin(priorInfo, histinfo$density)
    if(plot){
      rect(breaks[-length(breaks)], rep(0, length(breaks) - 1),
        breaks[-1], minHt, col = "green", border = "white")
    }
    overlap <- sum(minHt * diff(histinfo$breaks))
    if (is.function(prior)){
        if(plot){
          lines(histinfo$mids, priorInfo, lwd = 2, col = "brown")
          text(mean(breaks), 0, paste0("overlap = ", round(overlap * 100),
                                       "%"), pos = 3, cex = cex)
        }
    }
    return(overlap)
}

#############################################################################
##############################
#############################################################################
#############################
```

7. Session information

```
options(width=80)
sessionInfo()

R version 4.5.2 (2025-10-31)
Platform: x86_64-apple-darwin20
Running under: macOS Tahoe 26.3.1

Matrix products: default
BLAS:   /Library/Frameworks/R.framework/Versions/4.5-x86_64/Resources/lib/lib
Rblas.0.dylib
LAPACK: /Library/Frameworks/R.framework/Versions/4.5-x86_64/Resources/lib/lib
Rlapack.dylib;  LAPACK version 3.12.1

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

time zone: America/Vancouver
tzcode source: internal

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
```

Appendix B

```
 [1] tidyr_1.3.2      tibble_3.3.1     R2jags_0.8-9     rjags_4-17
 [5] purrr_1.2.1      plyr_1.8.9       kableExtra_1.4.0 HDInterval_0.2.4
 [9] ggplot2_4.0.2    dplyr_1.2.0      flextable_0.9.11 coda_0.19-4.1

loaded via a namespace (and not attached):
 [1] generics_0.1.4          fontLiberation_0.1.0    xml2_1.5.2
 [4] stringi_1.8.7           lattice_0.22-9          digest_0.6.39
 [7] magrittr_2.0.4          evaluate_1.0.5          grid_4.5.2
[10] RColorBrewer_1.1-3      fastmap_1.2.0           jsonlite_2.0.0
[13] zip_2.3.3               viridisLite_0.4.3       scales_1.4.0
[16] fontBitstreamVera_0.1.1 R2WinBUGS_2.1-24        textshaping_1.0.5
[19] abind_1.4-8             cli_3.6.5               rlang_1.1.7
[22] fontquiver_0.2.1        withr_3.0.2             yaml_2.3.12
[25] otel_0.2.0              gdtools_0.5.0           parallel_4.5.2
[28] tools_4.5.2             officer_0.7.3           uuid_1.2-2
[31] boot_1.3-32             vctrs_0.7.2             R6_2.6.1
[34] lifecycle_1.0.5         stringr_1.6.0           ragg_1.5.2
[37] pkgconfig_2.0.3         pillar_1.11.1           gtable_0.3.6
[40] data.table_1.18.2.1     glue_1.8.0              Rcpp_1.1.1
[43] systemfonts_1.3.2       xfun_0.57               tidyselect_1.2.1
[46] rstudioapi_0.18.0       knitr_1.51              farver_2.1.2
[49] htmltools_0.5.9         patchwork_1.3.2         labeling_0.4.3
[52] svglite_2.2.2           rmarkdown_2.31          compiler_4.5.2
[55] S7_0.2.1                askpass_1.2.1           openssl_2.3.5
```