# Using Integrated Models to Improve Management of Imperiled Salmon and Steelhead

Thomas Ward Buehrens

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2024

Reading Committee:
Daniel E. Schindler, Chair
André E. Punt
Martin C. Liermann

Program Authorized to Offer Degree:
School of Aquatic and Fishery Sciences

University of Washington

# Abstract

## Using Integrated Models to Improve Management of Imperiled Salmon and Steelhead

Thomas Ward Buehrens

Chair of the Supervisory Committee

Daniel E. Schindler

School of Aquatic and Fishery Sciences

Exploited wild animals include some of the most charismatic and celebrated species globally. They feed civilizations, are foundational to economies, are the subject of socially and culturally important recreational and ceremonial harvests and provide ecosystem services beyond human use. Yet, despite their importance to humanity, the history of management of these species is checkered, demonstrating both successes involving stable populations and long-term sustained harvests, and failures including overexploited species where population collapses and even extinction have occurred.

Sustainable exploitation depends entirely on the fundamental principle of compensatory recruitment, wherein population growth rates approach zero when populations are close to their carrying capacities, but as abundance declines due to harvesting, per capita growth rates increase in response to decreasing competition, thereby enabling populations to recover toward their

carrying capacities. However, to ensure harvest is implemented sustainably, answers to a few fundamental questions are required: 1) how many animals does a population contain? How many are being harvested? and, 3) What is the relationship between adult abundance and recruitment? Although natural resource managers have attempted to answer these questions for decades, the need for improved answers is greater than ever for species like Pacific Salmon (*Oncorhynchus* spp.) in the continental United States, where most populations are listed under the Endangered Species Act and face elevated risk of extinction yet remain subject to exploitation. This dissertation uses modern statistical modeling to address each of the fundamental questions necessary for the sustainable management of exploited but imperiled Pacific Salmon in Washington State.

Chapter 1 estimates the relationship between spawner abundance and recruitment to the smolt life stage for coastal steelhead (*O. mykiss*) populations in Washington State. By re-parameterizing the Beverton-Holt and hockey-stick stock-recruitment functions, the study incorporated habitat characteristics to explain variation in biological reference points among populations. The findings indicated strong effects of habitat quantity and quality on smolt carrying capacity, with estimates of capacity and the spawner abundance producing the maximum sustained yield ($S_{MSY}$) aligning with previous studies but demonstrating previously underappreciated sensitivity to smolt and kelt survival. The study highlighted the importance of geomorphic features in informing estimates of smolt capacity in stream-rearing salmonids and identified a minimum smolt-to-adult return rate required for population viability at around two percent based on maximum per-capita smolt productivities around 50 in the least productive populations.

Chapter 2 focuses on estimating the abundance of coho salmon (*O. kisutch)* for Washington State's populations in the Lower Columbia Evolutionarily Significant Unit (ESU). The study used data from an ESU-wide monitoring program initiated in 2010, which included spawning ground surveys, mark-recapture experiments, and trap-and-haul counts of migrating coho at dams, complemented by fishery catch estimation programs. The multivariate state-space integrated population model that was developed provided estimates of coho salmon population parameters, revealing that coho salmon abundance ranged between 12-86,000 wild spawners (median = 28,310) from 2010-2022. Several populations periodically exceeded their ESA recovery goals but remained below densities seen in healthy populations elsewhere.

Chapter 3 develops a Bayesian multivariate state-space model to improve estimates of catch in sport fisheries using creel survey and effort count data. This model, applied to data from the 2021 Skagit River wild winter steelhead catch-and-release sport fishery, robustly quantified uncertainty and was developed to flexibly accommodate multiple study designs. The model estimates of catch were used within a probabilistic management framework, helping fishery managers balance the risks of overfishing and unnecessarily restricting fishing opportunities. This approach demonstrated how robust quantification of uncertainty could optimize study designs for desired precision and cost efficiency.

## Acknowledgements

I thank my advisor, Daniel Schindler, for being everything and more that a student could ever hope for in an advisor. Daniel's passion for ecology, his question-driven approach to science, his methodological dexterity and his ability to dive into new topics and ask questions previously unidentified, are without comparison. Beyond that, I couldn't have had a stronger cheerleader and a more supportive and thoughtful person to mentor me through the trials and tribulations of doctoral studies.

I thank my committee for challenging me, forcing me to think about my work from different perspectives, and for their uncommonly and selflessly prompt responses to my every need. André Punt made me love and feel confident in math and programming in ways I never imagined and inspired me with his work ethic and his ironclad devotion to the success of his students and the School of Aquatic and Fishery Sciences. Martin Liermann straddled the chasm between quantitative tools and salmon ecology and was the most patient of mentors when I couldn't grasp simple concepts. Mara Zimmerman not only challenged me to think about the management and conservation applications of my work, but as my supervisor at WDFW for the first few years of my PhD studies, granted me to the opportunity to go back to school while working full time. Christian Torgersen was an incredibly supportive Graduate School Representative and brought helpful perspectives on landscape ecology to my work. Mark Scheuerell was my latest committee addition but stepped up to help me get through the finish line in a way that was very fitting given the less formal mentorship he had offered throughout my studies.

I thank my "unofficial committee" including two of my oldest and most important mentors. George Pess has spent the better part of two decades helping me be a better scientist

and perhaps even more importantly, a better person. I have learned a lot from his kindness and his generosity with his time. Dan Rawding inspired me to take the job I have at WDFW and schooled me in the ways of salmon abundance and survival estimation. His passion for the science and the management of salmon and steelhead in Washington have left an indelible mark that the fish and generations of anglers will be grateful for. Thanks also to Tom Quinn, Peter Kiffney, and Todd Seamons, whom I've learned a great deal from, greatly respect, and helped make me the scientist I am today starting from my earliest days in my first tour of duty in grad school.

I thank the Schindler Lab for being the most amazing collection of smart, passionate, hardworking, and helpful people one can find. As important as advisors are, graduate students learn as much or more from their lab-mates. I appreciated endless hours of statistical whiteboard sessions with Lukas DeFilippo and Jan Ohlberger. Sean Brennan was an amazing friend and mentor who was taken from us too soon and is dearly missed. Adrianne Smits, Michaela Lowe, Sarah O'Neil, and Ashley Townes were always there to pick me up when I was down or wanted to talk about something other than math.

I thank the Washington Department of Fish and Wildlife for supporting me to go back to school. Kelly Cunningham authorized my studies and always supported me, as did Bryce Glaser, who trusted that I would still support Lower Columbia region needs despite moving Seattle. Kale Bentley, Charlie Cochran, and Jeremy Wilson tolerated me as an occasionally errant supervisor and supported me to the end. Neala Kendall helped my confidence as a scientist and gave me new appreciation for what collaboration can be between colleagues with complementary skillsets. Laurie Peterson was a tireless cheerleader for me.

I thank my family and friends for never giving up hope that I would one day finish my dissertation, even if in the meantime, they cracked many jokes. I thank my father, Paul, and my grandfather, Cliff, for taking me fishing and lighting the fire that got me here. I thank my mother, Tish, for instilling the conservation ethic and a passion for wild places and the outdoors. I thank my brother for teaching me what perseverance can achieve and how to use underdog status to make you stronger. I thank my oldest fishing buddy, Will Atlas, for his personal and professional friendship tracing back to elementary school. I thank Bill and John McMillan for teaching me to be a better naturalist, and for their love for wild rivers and wild fish. Finally, I thank my wife, Allyson. You are my best friend and my biggest fan. I admire you more than you know, and I am the luckiest man on earth to call you my wife. Thank you, one final time, for your endless patience while I dawdled along, and for always believing in me.

**Table of Contents**

# Chapter 1: Estimates of biological reference points for coastal steelhead populations in Washington State

## 1.A. Abstract

Compensatory recruitment is a critical process that enables the sustainable exploitation of species and reduces extinction risk. Understanding the role of density dependence on recruitment is therefore important yet can be challenging due to recruitment variation resulting from other factors. In stream-rearing anadromous salmonids such as steelhead trout (*Oncorhynchus mykiss*) density dependence is particularly strong during freshwater residency, which simplifies efforts to estimate vital rates needed to establish biological reference points. However, because juvenile abundance data are often limited, hierarchical stock-recruitment models provide a powerful means to establish biological reference points by sharing data among similar populations. The inclusion of covariates that describe landscape characteristics can explain variation in reference points among populations as well as enable more precise predictions for populations without data, further extending the utility of hierarchical stock- recruitment models. We re-parameterized the Beverton-Holt and hockey-stick stock-recruitment functions so that carrying capacity was a function of the quantity and quality of stream habitat, which was described by a principal component axis representing watershed attributes. Both habitat quantity and quality had strong effects on carrying capacity. Our estimates of smolt capacity and spawner abundance producing the maximum sustained yield $S_{MSY}$ were similar to estimates developed in previous studies, although $S_{MSY}$ was sensitive to the smolt-to-adult and kelt (survival between spawning events) survival rates. Based on our estimates of smolt productivity, the minimum smolt-to-adult return rate required for population viability appears to be around two percent, which is similar to

estimates from previous studies. Our modeling approach illustrates the utility of geomorphic features in informing estimates of smolt capacity in stream-rearing salmonids.

## 1.B. Introduction

Estimates of population demographic parameters and biological reference points are essential for assessing population viability and for managing exploited species to achieve optimal yields. For all exploited species, including fishes like semelparous Pacific Salmon (*Oncorhynchus* spp.), parameters including the maximum per capita recruitment rate (i.e., productivity) and the maximum asymptotic population recruitment (i.e., capacity) are important in determining the equilibrium unexploited population size, fishery exploitation rates that maximize sustainable yields ($F_{MSY}$), and the population abundances that produce those yields (e.g., $S_{MSY}$) (Hilborn and Walters 1992). These parameters, in conjunction with estimates of variation in demographic rates, may also be used to project population dynamics into the future to assess viability and the risk of extinction (Buhle et al. 2018). However, common problems that inhibit the estimation of population demographic parameters for Pacific Salmon populations include insufficient data and the noise-generating effects of density-independent environmental variability (Schnute and Kronlund 2002) that obscure the relationship between density and survival.

For stream-rearing anadromous salmonids, the strength of density dependence in freshwater prior to ocean entry provides a means to eliminate much of the noise in population data to estimate demographic parameters. Spawner-recruit analyses using the abundance of recruits prior to ocean entry may remove considerable density-independent noise generated in the marine phase of the life cycle, revealing the density-dependent relationship in the freshwater phase of the life cycle (Milner 2003, Bradford et al. 2000). Steelhead, *Oncorhynchus mykiss*, the anadromous form of rainbow trout (hereafter, steelhead), have one of the most protracted freshwater rearing phases (1-5 years) among anadromous salmonids (Quinn 2005). In the freshwater environment, this species can occur in high densities and establishes territories (e.g., Keeley 2000), resulting in density-dependent survival (Ward and Slaney 1993). Consequently, stock-recruitment models using the abundance of smolt recruits may ease the estimation of density-dependent recruitment relationships for steelhead relative to models using the abundance of adult recruits.

In contrast, steelhead occur in low densities in the marine environment and resources are shared among many populations in common rearing areas, meaning the abundance of any individual population negligibly affects aggregate abundance and density-dependent survival of steelhead recruits at sea. Therefore, like Atlantic salmon (e.g., Jonsson et al, 1988), survival of individuals from a particular population during the marine phase may be relatively independent of the abundance of that population. The obscuring effect of density-independent marine survival on stock-recruitment relationships for steelhead is exacerbated by considerable variability in marine survival, which may fluctuate by more than an order or magnitude at decadal or shorter time scales (Ward 2000). Additionally, previous work suggests that recruitment variation for salmonids is correlated among stocks at large spatial and temporal scales, which is likely a result of spatial and temporal coherence in marine survival (e.g., Myers et al. 1997; Peterman et al. 1998; Zimmerman et al. 2015) and complicates estimation of parameters of stock-recruitment relationships because the parameters for some recruitment functions are sensitive to marine survival (Ward 2000). In contrast, there may be less spatial correlation in freshwater recruitment variation among years or even within a year for populations from the same region (Bradford 1999).

The limiting nature of freshwater habitat on the population dynamics of steelhead has been long-appreciated and has previously been incorporated into the estimation of management reference points (Gibbons et al. 1985; Johnson et al. 2000; Tautz et al. 1992). In Washington State, Gibbons et al. (1985) estimated juvenile capacity in western Washington rivers by enumerating juvenile steelhead as parr during the late summer low-flow period (when freshwater habitat was assumed to be most limiting) and prior to assumed ocean entry by the juveniles the next spring. Parr densities per unit area of stream were calculated for four strata based on stream gradient and these densities were then used to extrapolate potential parr capacity to un-surveyed reaches. Short time-series (3-6 years) of spawner and adult recruit data were available from five watersheds and these abundance data were standardized among watersheds per unit of estimated juvenile steelhead parr capacity and combined into a single stock-recruit analysis. A Beverton-Holt model was fit to these data allowing for estimates of $S_{MSY}$ per unit of juvenile capacity, which could then be used to calculate $S_{MSY}$ in different watersheds. Their approach allowed management reference points to be calculated in watersheds where data were not available with the caveats that the influence of marine survival was ignored and the effects of iteroparity were omitted. Reference points (e.g., $S_{MSY}$) provided by this method have been used to manage fisheries in much of coastal Washington for over three decades. Similar habitat-based methods were used to estimate recovery thresholds and population viability related to the listing of Puget Sound steelhead under the USA Endangered Species Act (Hard et al. 2013). However, the extrapolated estimates of population-scale juvenile capacity in both modeling efforts were based on limited reach-scale observations rather than population-scale estimates of juvenile recruitment. New adult and juvenile abundance data have been collected and considerable advances have been made in the analytical methods used to estimate population reference points in the four decades since original reference points were developed, enabling a re-assessment of reference points for coastal steelhead populations that includes the influence of marine survival and iteroparity.

Hierarchical spawner-recruit models provide an effective way to estimate population productivity and capacity in salmonids, particularly where data are limited or noisy. In a hierarchical framework, the parameters describing population-specific models are shared among populations, and parameter values for data-poor locations are objectively informed by sites with high quality data. Such methods have been used to estimate population parameters for Chinook salmon (*O. tshawytscha*; Liermann et al., 2010), coho salmon (*O. kisutch;* Barrowman et al. 2003), and Atlantic salmon (*Salmon salar;* Michielsens and McAllister 2004). This approach provides empirical estimates of population productivity and capacity and can be conducted with adult spawners and their juvenile (smolt) recruits (e.g., Michielsens and McAllister 2004), thus eliminating the portion of noise in typical stock-recruitment datasets caused by density-independent recruitment variation associated with smolt to adult survival.

A critical assumption of hierarchical spawner-recruit analyses is that hierarchically modeled parameters (e.g., productivity, capacity) are drawn from a hyper-distribution with a common central tendency and variance among populations. This assumption may be relatively robust with respect to productivity but is clearly violated or requires an excessively large hyper-variance with respect to capacity when analyses are based populations belonging to watersheds of differing size and habitat quality (Bradford et al. 2000; Sharma and Hilborn 2001). Liermann et al. (2010) dealt with this assumption by accounting for watershed size in their model structure to estimate Chinook salmon reference points; however, their analysis did not incorporate other watershed characteristics that could influence the capacity per unit of habitat quantity. Using juvenile recruits in such a stock-recruitment analysis and modeling capacity as function of habitat area and covariates related to habitat quality provides a direct means of testing and validating habitat area- and quality-based estimates of juvenile carrying capacity.

Our objective was to estimate productivity, capacity and $S_{MSY}$ for coastal steelhead populations in Washington State. We performed a hierarchical Bayesian state-space stock-recruitment analysis on spawner and smolt abundance and age data for 14 coastal steelhead populations. We accounted for systematic differences in capacity among populations resulting from differences in habitat quantity and quality by modeling capacity as function of accessible stream length and the geomorphic features of watersheds that likely affect habitat quality. Finally, we evaluated the sensitivity of estimates of $S_{MSY}$ to variation in smolt-to-adult and post-spawning (i.e., kelt) steelhead survival.

## 1.C. Methods

### *Spawner Data*

Spatially paired estimates of steelhead spawner and smolt abundance were available from 14 coastal (i.e., west of Cascade Mountain crest) winter and summer run steelhead populations in Washington State (Table 1, Figure 1). Spawner data for each population varied in length between two and 38 years, with 12 datasets greater than 10 years in length. Steelhead spawner abundance estimates (Table 1) were made using a variety of methods including census counts at weirs and dams, Petersen mark-recapture estimators, and redd count expansions (e.g., Gallagher et al., 2007). Year and population-specific observation error estimates were available for many spawner estimates based on mark-recapture methods, whereas census spawner abundance estimates were given an observation error CV of 0.01 because these abundances were assumed to be measured nearly without error. For population-by-year combinations where empirical sampling-based estimates of observation error were not available, observation error was estimated within the model.

For some populations, hatchery-origin steelhead spawners were present, and, in some cases direct estimates of their abundance were available. In other cases, their abundance was estimated by multiplying the number of smolts released by year- and population-specific estimates of marine survival and assumed harvest and hatchery broodstock collection rates. In comparison to wild spawners, hatchery steelhead spawning in the wild are known to have lower relative reproductive success (RRS) than their wild counterparts (Araki et al., 2008), and it was therefore necessary to adjust these hatchery spawners into “wild spawner equivalents”. This was done by multiplying their abundance by RRS values calculated to the smolt stage that were reported in the literature. An RRS value of 0.3 was used for hatchery steelhead derived from heavily domesticated “segregated” broodstocks (range 0.18-0.37 RRS; Kostow et al. 2003), whereas a value of 0.85 selected for hatchery steelhead derived from “integrated” broodstocks that wholly or partially use wild spawners as their broodstocks. This value was intermediate between estimates of RRS to the adult stage for steelhead and to smolt stage for coho salmon as reviewed by Araki et al. (2008).

### *Smolt Data*

Wild smolt abundance estimates were made for each calendar year using either census counts or pooled or time-stratified Petersen mark-recapture estimators (e.g., Darroch 1961, Bonner and Schwarz 2009) based on smolts captured at rotary screw traps (e.g., Volkhardt et al. 2007). Datasets varied in length between one and 36 years, with nine datasets greater than ten years in length. Observation error CVs were generally available for most mark-recapture smolt abundance estimates. For mark-recapture smolt abundance estimates that did not have observation error estimates, the observation error CV was assigned the highest value in the time series for a particular population. Census smolt counts were assigned an observation error CV of 0.01.

Steelhead most commonly smolt at ages two and three in western Washington, however, a portion of individuals smolt at ages one and four and age composition varies among populations. Representative collections of scales were available for some populations and years and enabled estimation of migration year age proportions which could then be used to assign smolts in a migration year as recruits from their respective parental brood years. Smolt abundance estimates were only made for ocean-migrating individuals during spring months. For traps located directly above saltwater, this essentially represented the entire freshwater production from the upstream watershed, since steelhead entering marine waters in months other than the spring, or at life stages other than smolt, are known to have low survival (Hall et al. 2016) and are relatively rare in these ecosystems. Although steelhead may move past these locations at multiple life stages, we only included smolts, which were identified by their size, external coloration, and morphology, in order to estimate the density-dependent production of this life stage based on habitat upstream of traps in a comparable way among all trap sites.

### *Recruitment Functions*

We selected hockey-stick (eq. 1; Barrowman and Myers 2000) and Beverton-Holt (eq. 2; Beverton and Holt 1957) stock-recruit models to estimate the relationship between spawner abundance and the recruitment of smolt offspring. The hockey stick model follows the form:

$$R = \min\ (\alpha S,\ \alpha S^* ) = \begin{cases} \alpha S, & S < S^* \\ \alpha S^*, & S \geq S^* \end{cases} \tag{1}$$

where $R$, the number of recruits, is the product of the productivity, $\alpha$, and the spawner abundance $S$, when $S$ is less than the abundance needed to fully seed the rearing habitat; $S^*$. When $S \geq S^*$, recruitment is at carrying capacity, and is therefore constant and equal to $\alpha S^*$. The hockey-stick function results in a break-point linear function shaped like a hockey-stick, which has attractive properties for modeling recruitment of stream rearing salmonids. First, estimates of maximum per-capita recruitment from this model are lower than those produced by the Ricker and Beverton-Holt models (Barrowman and Myers 2000), both of which tend to over-estimate per capita recruitment at low spawner abundance (Myers et al. 1994) potentially resulting in overestimates of sustainable exploitation rates and underestimates of risk of extinction. Additionally, the shape of the hockey stick curve may be particularly appropriate for territorial fishes such as steelhead, for which density-dependent mortality may be relatively minor until all territories are occupied (e.g., coho salmon; Bradford et al. 1997), resulting in relatively constant per-capita recruitment up to this threshold.

We parameterized the Beverton-Holt model as:

$$R = \frac{\alpha S}{\left(1 + \frac{\alpha S}{\beta}\right)} \tag{2}$$

where $R$, the number of recruits, is the product of the productivity, $\alpha$, and the spawner abundance, $S$, divided by a saturating function describing the magnitude of density-dependence where $\beta$ is the "habitat capacity" for recruits. Habitat capacity, which is the maximum expected recruitment, is greater than the carrying capacity, the expected recruitment at unfished equilibrium spawner abundance. The Beverton-Holt model is a useful alternative to the hockey-stick model for territorial fishes, particularly if territoriality is imperfect, or if suboptimal territories or habitat patches become progressively occupied as cohort abundance increases. In this model, recruitment asymptotically increases even as spawner abundance surpasses replacement levels. Additionally, the Beverton-Holt model assumes continuously increasing per-capita recruitment as abundance declines, thereby providing a useful upper bound for productivity in territorial fishes.

Additional models that allow for overcompensation (e.g., Ricker) may be less appropriate for territorial fishes since territoriality allows for unequal resource sharing among individuals and thus the maintenance of high survival for those individuals which do successfully defend territories (Hilborn and Walters 1992). Such unequal resource-sharing and the maintenance of high survival for a subset of the population are not expected attributes of populations that exhibit overcompensation (i.e. declining recruitment because of strong density-dependence at higher spawner levels).

### ***Modifications to Traditional Univariate Stock-Recruitment Analysis***

We modeled stock-recruitment relationships for all populations simultaneously using hierarchical models, extending similar work done for coho salmon (Barrowman and Myers 2003), Atlantic salmon (Michielsens and McAllister 2004), and Chinook salmon (Liermann et al. 2010). Hierarchically modeling productivity among populations relies upon the assumption that the maximum per-capita recruitment of each population belongs to a hyper-distribution with a common mean and variance and population-specific productivities are random effects. Hierarchically modeling carrying capacity of populations belonging to watersheds of differing habitat quantity implies the need to incorporate variables accounting for differences in habitat quantity or quality, and presumably, carrying capacity (e.g., Liermann et al. 2010).

We re-parameterized both stock-recruitment models to model capacity as log-linear function of habitat covariates. First, we standardized capacity estimates for variation in habitat quantity among populations by estimating capacity as the product of capacity density (smolts per channel length) multiplied by the estimated channel length of occupied steelhead habitat with >= 50% probability of steelhead occurrence (see Eq. 10).  This approach has been used previously for coho salmon (Sharma and Hilborn 2001; Korman 2015, Bradford 1999, Barrowman and Myers 2003) and Chinook salmon (Liermann et al 2010). We further accounted for habitat effects on capacity by including geomorphic variation among rivers, characterized by channel attributes for every 100 m reach of occupied habitat above the smolt traps from TerrainWorks NetMapTools (Benda et al. 2007). We limited our candidate variables *a priori* to four variables thought to potentially influence habitat quality for rearing steelhead. This is because, although several dozen watershed attributes were available from TerrainWorks, with only 14 populations from which to estimate the effect of habitat covariates on among-population variability in capacity, we wanted to avoid competing multiple habitat covariate models given our limited statistical power to perform model selection. The selected covariates included: 1) the log of the ratio of mean annual discharge at the smolt trap to occupied habitat length, which provided a measure of how large channels were, 2) the mean of the logit transformed channel gradients in each 100 m reach, 3) the mean of log transformed channel widths in each reach, and 4) the logit-transformed cumulative watershed hillslope gradient at each smolt trap (Table 2). Each normalized variable was then standardized, and we performed a Principal Component Analysis (PCA) to reduce co-linearity in our covariate dataset and reduce its dimensionality (Table 2; Figure 2).

Finally, we fit our stock-recruit model in a state-space framework to accommodate partitioning of observation and process error (e.g., Fleischman et al. 2013), which can reduce bias in population parameters associated with aging error (Zabel and Levin 2002) and imprecise estimates of spawners and recruits (e.g., Ludwig and Walters 1981).

### *Process Model*

The logarithms of spawner abundance $S$ for each population *p* and year *t* were assumed to follow a log-normal random walk:

$$\log\left(S_{t,p}\right) \sim \text{Normal}(\log\left(S_{t-1,p}\right),\ \sigma_p^S) \tag{3}$$

These spawners produced smolt offspring $R$, the abundance of which was assumed to be lognormally distributed (to incorporate recruitment process error):

$$R_{t,p} \sim \text{Lognormal}\left(\mu^{R}_{t,p}, \sigma^{R}_{p}\right) \tag{4}$$

where $\mu^{R}$ was the expected log-recruitment from either the hockey stick model (eq. 1) substituting population-specific smolt capacity $K_p$ for $a_p S^{*}_{p}$, or the Beverton-Holt model (eq. 2), i.e.:

$$\mu^{R}_{t,p} = \begin{cases} \log\left(\alpha_p S_{t,p}\right), & \text{if } \log\left(\alpha_p S_{t,p}\right) < \log\left(K_p\right) \\ \log\left(K_p\right), & \text{if } \log\left(\alpha_p S_{t,p}\right) \geq \log\left(K_p\right) \end{cases} \tag{5}$$

$$\mu^{R}_{t,p} = \log\left[\frac{\alpha_p S_{t,p}}{\left(1+\frac{\alpha_p S_{t,p}}{\beta_p}\right)}\right] \tag{6}$$

Smolt recruits from brood year *t* were then multiplied by year- and population-specific brood year smolt age proportions $pBY_{t,p,a}$ (estimated parameters that were subject to a sum to one constraint each year) to produce the expected numbers of smolt recruits of each age *a* that out-migrated in a particular migration (calendar) year $MYsmolts_{t,p,a}$:

$$R_{t,p} \times pBY_{t,p,a} = MYsmolts_{t+a,p,a} \tag{7}$$

which were then summed by calendar year to estimate the total smolt outmigration (comprised of mixed parental brood years) in a particular out-migration year, and could each also then be divided by that sum to calculate the population-specific outmigration year-age proportions $pMY_{t,p,a}$:

$$pMY_{t,p,a} \frac{MYsmolts_{t,p,a}}{\sum_{a=1}^{4} MYsmolts_{t,p,a}\, TotMYsmolts_{t,p}} \tag{8}$$

The log of the population-specific smolt recruit capacity $\mu_p^K$ was calculated as a log-linear function of the PC1 of habitat features, including an offset for the length in kilometers of habitat occupied by steelhead $L_p$. The population-specific capacity $K_p$ was then distributed log-normally around $\mu_p^K$ with an among-population capacity parameter standard deviation $\sigma^K$:

$$\mu_p^K = \log(L_p) + \beta_0 + \beta_1 PC1 \tag{9}$$

$$K_p \sim \text{LogNormal}(\mu_p^K, \sigma^k) \tag{10}$$

For the Beverton-Holt model, $\beta_p$ and $\mu_p^\beta$ were substituted for $K_p$ and $\mu_p^K$ in equations 9 and 10.

### *Observation Model*

A lognormal likelihood was used to relate estimates of spawner abundance to the unobserved latent state estimates of spawner abundance $S_{t,p}$:

$$E_{t,p} \sim \text{LogNormal}(\log(S_{t,p}), \sigma_{t,p}^{obs}) \tag{11}$$

where the observed point estimate (assumed to be related to a parameter median) of spawner escapement $E_{t,p}^{obs}$, had a year- and population-specific observation standard error, $\sigma_{t,p}^{obs}$. Because observation standard deviations for spawner abundance were only available for some years, we assumed the population and year-specific observation error standard deviations were in turn drawn from lognormal distributions (eq. 12) with a population-specific mean observation error standard deviation $\mu_p^{\sigma^{obs}}$ and year-to-year variability in observation error standard deviations $\sigma^{\sigma^{obs}}$ that was assumed to be the same among populations, which allowed the model to impute values for population and year combinations without values:

$$\sigma_{t,p}^{obs} \sim \text{LogNormal}(\log{(\mu_p^{\sigma^{obs}})}, \sigma^{\sigma^{obs}}) \tag{12}$$

This hierarchical structure enabled imputation of observation error standard deviation estimates for years for which estimates were missing for particular populations.

An additional lognormal likelihood related the observed total smolt out-migrant abundance in a particular calendar year, $EstMYsmolts_{t,p}$ to the true smolt abundance in that migration year $TotMYsmolts_{t,p}$ (a latent unobserved state variable described in the process model), again with year and population-specific observation error standard deviations $\sigma_{t,p}^{obs}$, which were available for all year by population combinations:

$$EstMYsmolts_{t,p} \sim \text{LogNormal}\left(\log\ (TotMYsmolts_{t,p}), \sigma_{t,p}^{obs}\right) \tag{83}$$

A multinomial likelihood related the out-migration-year and population-specific subset of scale-aged smolts $ScaleAgeCount_{t,p,a}$ to the population-specific outmigration year age proportions of the recruits $pMY_{t,p,a}$(a latent unobserved state vector):

$$ScaleAgeCount_{t,p,a} \sim \text{Multinomial}(pMY_{t,p,a}, \sum_{a=1}^{4} ScaleAgeCount_{t,p,a}) \tag{14}$$

***Priors***

Spawner abundance for the first adult abundance state was given a vague lognormal prior that minimally constrained spawner abundance in the first year:

$$S_{t,p} \sim \text{LogNormal}(0, 8) \tag{15}$$

The brood year smolt recruit age proportions were estimated in a three-level logit-normal hierarchy. First, global log odds ratio components $\mu_a$ were estimated for ages two through four with the first log odds ratio component $\mu_{a=1}$ arbitrarily fixed to zero:

$$\mu_{a>1} \sim \text{Normal}\,(0, 5) \tag{16}$$

Population-specific and year-specific within-population adjustments to these global log-odds ratio components were then assumed to be distributed normally:

$$\delta_{p,a}^{pop} \sim \text{Normal}\,(0, \sigma^{pop}) \tag{17}$$

$$\delta_{t,p,a}^{year} \sim \text{Normal}\,(0, \sigma^{year}) \tag{18}$$

The global (eq. 19), population-specific (eq. 20), and population and year-specific (eq. 21), brood year smolt recruit age proportions were then calculated via exponentiation of the log odds ratios:

$$pBY_a = \frac{\exp\,(\mu_a)}{\sum_{a=2}^{a=4} \exp(\mu_a) \; + \; \exp\,(0)} \tag{19}$$

$$pBY_{p,a} = \frac{\exp\,(\mu_a + \; \delta_{p,a}^{pop})}{\sum_{a=2}^{a=4} \exp\left(\mu_a + \; \delta_{p,a}^{pop}\right) \; + \; \exp\,(0 \; + \; \delta_{p,a}^{pop})} \tag{20}$$

$$pBY_{t,p,a} = \frac{\exp\,(\mu_a + \; \delta_{p,a}^{pop} + \; \alpha_{t,p,a})}{\sum_{a=2}^{a=4} \exp\left(\mu_a + \; \delta_{p,a}^{pop} + \; \delta_{p,a}^{year}\right) \; + \; \exp\,(0 \; + \; \delta_{p,a}^{pop} + \; \delta_{p,a}^{year})} \tag{1}$$

The population-level and year-within-population level age composition error standard deviations were given vague priors:

$$\sigma^{pop} \sim \text{half} - \text{Cauchy}\,(0,1) \tag{92}$$

$$\sigma^{year} \sim \text{half} - \text{Cauchy}\,(0,1) \tag{23}$$

Population-specific productivity $\alpha_p$ was modeled hierarchically on the log scale as a normal distribution based on the among-population mean $\mu_\alpha$ and standard deviation $\sigma_\alpha$ of productivity:

$$\alpha_p \sim \text{LogNormal}\,(\mu_\alpha, \; \sigma_\alpha) \tag{24}$$

The among-population mean productivity $\mu_\alpha$ was estimated with a vague prior that constrained it to biologically plausible values (eq. 26), and the among population variability was given a vague prior as well (eq. 27):

$$\mu_\alpha \sim \text{Normal}(\log(55), 1.5) \tag{25}$$

$$\sigma_\alpha \sim \text{Cauchy}(0,1) \tag{26}$$

Standard deviations for process errors in smolt recruitment $\sigma_p^R$ and spawner abundance $\sigma_p^S$ were modeled hierarchically among populations in log space:

$$\log(\sigma_p^R) \sim \text{N}(\log(\mu^{\sigma^R}), \sigma^{\sigma^R}) \tag{27}$$

$$\log(\sigma_p^S) \sim \text{N}(\log(\mu^{\sigma^S}), \sigma^{\sigma^S}) \tag{28}$$

and the among-population process error means ($\mu^{\sigma^R}$and $\mu^{\sigma^S}$) and standard deviations ($\sigma^{\sigma^R}$ and $\sigma^{\sigma^S}$) were given vague priors:

$$\mu^{\sigma^R} \sim half - \text{Cauchy}(0,1) \tag{2910}$$

$$\mu^{\sigma^S} \sim half - \text{Cauchy}(0,1) \tag{110}$$

$$\sigma^{\sigma^R} \sim half - \text{Cauchy}(0,1) \tag{31}$$

$$\sigma^{\sigma^S} \sim half - \text{Cauchy}(0,1) \tag{122}$$

The population-specific mean observation error standard deviations $\mu_p^{\sigma^{obs}}$ and the global among-year standard deviation in spawner observation errors, were given also vague priors:

$$\mu_p^{\sigma^{obs}} \sim \text{half} - \text{Cauchy}(0, 1) \tag{133}$$

$$\sigma^{\sigma^{obs}} \sim \text{half} - \text{Cauchy}(0, 1) \tag{144}$$

Finally, the among population capacity-habitat regression intercept $\beta_0$, standardized coefficient $\beta_1$ for the effect of PC1, and population-specific capacity residual error standard deviation $\sigma^k$ were all given vague priors:

$$\sigma^k \sim half - \text{Cauchy}(0,1) \tag{155}$$

$$\beta_0 \sim \text{N}(0, 10) \tag{166}$$

$$\beta_1 \sim \text{N}(0,10) \tag{177}$$

***Calculating $S_{MSY}$***

Brooks et al. (2010) provided an equation to calculate the spawner abundance $S_{MSY}$ at which yield (in numbers of fish rather than weight) is maximized. We re-parameterized equation 4 from Brooks et al. (2010) to match the parameterization of the Beverton-Holt recruitment function we used and to adjust by smolt to maiden (first-time-spawner) adult return rate *SAR*, since the recruits in our model had not reached adulthood. The modified equation was:

$$S_{MSY} = \frac{\beta(\sqrt{\alpha \cdot SAR \cdot \phi_0} - 1)}{\alpha} \tag{188}$$

This equation provides identical estimates of $S_{MSY}$ to the estimate for semelparous fishes (e.g., salmon) provided by Hilborn and Walters (1992; Table 7.2) when the ratio of spawners to recruits at unfished equilibrium $\phi_0$ is 1.0. Brooks et al. (2010) outline how to calculate $\phi_0$ in units of eggs or spawners $S_0$ to recruits $R_0$ at unfished equilibrium (Eq. 2; Brooks et al. 2010):

$$\phi_0 = \frac{S_0}{R_0} = \sum_{a=r}^{A} \mu_a E_a \prod_{a=r}^{a-1} e^{-M_a} \tag{39}$$

where $A$ is the maximum of age $a$, $r$ is the age of recruitment, $\mu_a$ is the maturation probability by age, $E_a$ is the age-specific fecundity (eggs), and is used if $\phi_0$ is desired to be expressed in terms of eggs (it is otherwise set to 1), and $e^{-M_a}$ is the age-specific survival between spawning events. Adult age composition and fecundity-at-age data that could be used to estimate these parameters via a life-cycle model, were either inconsistently collected, or missing entirely for most of our populations. However, when $\mu_a$is set to zero for ages below the youngest reproductive age (e.g., typically 3-4 for female steelhead), $E_a$is fixed to one, and $e^{-M_a}$ is constant, $\phi_0$ may be approximated by:

$$\phi_0 = 1 + \frac{e^{-M}}{1 - e^{-M}} \tag{190}$$

Since we were missing data to estimate population-specific $\phi_0$, and thus iteroparous $S_{MSY}$, we instead chose to estimate $S_{MSY}$ assuming $\phi_0 = 1$ (semelparity), and then assess the sensitivity of these estimates relative to the true value of $S_{MSY}$ when $\phi_0 > 1$.

The ratio of "true" $S_{MSY}$ to semelparous $S_{MSY}$ calculated using the approximation in eq. 41 of $\phi_0$ is approximate because it assumes, absent the effects of fishing mortality, the same sex ratio and per capita reproductive output between maiden and repeat spawners, and constant adult survival between age classes. Although assuming age-specific adult mortality and fecundity are constant are not necessarily reasonable assumptions for steelhead, using equation 41 as parametrized enables an improved approximation of $S_{MSY}$relative to simply assuming steelhead are semelparous when lacking the rich demographic data needed for to calculate $S_{MSY}$ using the full age-structured model proposed by Brooks et al. (2010). A similar sensitivity analysis was unnecessary for hockey stick model because sustainable yield is always maximized when spawner abundance = $S^*$ (Eq. 2), which is invariant to smolt-to-adult survival and $e^{-M}$ because of the broken-stick shape of the recruitment curve; recruitment is always maximized at the spawner abundance where the "joint" in the regression occurs.

### *Fitting Models*

We fit our hierarchical stock-recruitment models in a Bayesian framework using Markov Chain Monte Carlo (MCMC) simulations to sample the posterior probability density functions for each parameter. We conducted our MCMC simulations using JAGS (Plummer 2003), which was operated in the R statistical programming environment (R Development Core Team 2014) using the R2jags package (Su and Yajima 2015). We obtained 400,000 MCMC samples from four chains, discarding the first 100,000 samples from each chain as burn-in and thinning to retain one sample in every 100, leaving 12,000 samples of the posterior distribution. We assessed chain convergence and stability by visually assessing trace-plots, calculating the Brooks-Gelman-Rubin statistic, and calculating the effective sample size (ESS; Gelman et al. 2013). All BGR statistics were < 1.05 and most were less than 1.01. Effective sample sizes were mostly in the thousands, however a small number of parameters had ESS of 100-1000.

## 1.D. Results

### *PCA of Habitat Attributes*

The PCA performed on four habitat attributes yielded orthogonal PC axes, the first of which explained 50% of the variation among watersheds (Figure 2). The first PC axis was most heavily influenced by watershed slope ($r = 0.66$), followed by the ratio of mean annual discharge

(MAD) to accessible channel length ($r = 0.56$), channel gradient ($r = 0.38$), and channel width ($r = 0.32$).

### ***Estimates of the Recruitment Relationship***

Fully Bayesian estimates of the recruitment relationship were generated for both Beverton-Holt (Figure 3) and hockey stick functions (Figure S1). This was done by plotting predicted recruitment against spawner abundance on a draw-by-draw basis across a range of spawner abundances. For the hockey stick model this resulted in smoothing of the median recruitment curve due to negative covariance between productivity and capacity. To evaluate how well the model would perform making predictions for unmonitored populations we produced both predictions for each population using population-specific random effects for productivity and capacity (blue lines) as well as using the random effect hyper distributions with means of zero (black lines; Figure 3, Figure S1).

### ***Smolt Capacity per Kilometer***

Population-specific capacity estimates ranged from 22 (HS model) and 26 (BH model) smolts per kilometer for Bingham Creek to 315 (HS model) and 427 (BH model) smolts per kilometer for the NF Toutle River (BH: Figure 4b, HS: Figure S2b; Table 3). There was a strong positive association between the standardized coefficient for the effect of PC1 ($\beta_1$) and estimated smolt capacity; HS median = 0.50 (95% CI = 0.30-0.70) and BH median = 0.54 (95%CI = 0.31-0.77) (Figure 4a, Figure S2a). Although PC1 explained a large portion of the variance in capacity estimates per kilometer among populations, there remained a considerable amount of uncertainty in posterior predictive estimates of this parameter (BH CV = 67%-72%, and HS CV = 55-60% depending on the quantile of PC1) (Figure 4a, Figure S2a, Table 3).

### *Smolt Productivity*

Population-specific estimates of productivity (maximum per-capita smolt recruitment) ranged from 44 to 75 and 81 to 112, respectively, for the hockey stick and Beverton-Holt models (Figure 5, Table 4). The posterior predictive distribution, displayed as the unknown population in Figure 5, had medians of 56 (80% CI = 35-91) and 97 (80% CI = 57-172) smolts per spawner for the hockey stick and Beverton-Holt models, respectively.

### *Spawner Density at MSY*

The posterior predictive distribution for the number of spawners per kilometer needed to produce the maximum sustainable yield varied substantially among watersheds, and, for the Beverton-Holt model, as a function of SAR and $e^{-M}$ (Figure 6; Table 4). For the hockey stick model $S_{MSY}$ occurred at median = 1.29 spawners per kilometer (80% CI 0.58-2.90) for low quality habitat (PC1 quantile = 0.10), median = 2.45 spawners per kilometer (80% CI 1.12-5.41) for average quality habitat (PC1 quantile = 0.50), and median = 4.65 spawners per kilometer (80% CI 2.09-10.52) for high quality habitat (PC1 quantile = 0.90) (Figure 6, Table 4).

Estimates of semelparous $S_{MSY}$ under the Beverton-Holt model when SAR = 2% ranged from median = 0.33 spawners per kilometer (80% CI 0.08-0.69) for low quality habitat to 1.31 spawners per kilometer (80% CI 0.32-2.78) for high productivity habitat (PC1 quantile = 0.90).

However, estimates scaled proportional to marine survival were substantially higher at 15% SAR; median = 2.61 spawners per kilometer (80% CI = 1.27-5.26) for low productivity habitat (PC1 quantile = 0.10) and 10.33 spawners per kilometer (80% CI 4.95-21.16) for high productivity habitat (PC1 quantile = 0.90). Beverton-Holt $S_{MSY}$ estimates were most similar to hockey-stick estimates of $S_{MSY}$ (which also equals $S$*) when SAR was equal to 5% (Figure 6; Table 5).

In addition, the negative bias of semelparous Beverton-Holt $S_{MSY}$ estimates relative to their true (iteroparous) counterparts was influenced both by SAR and by the survival rate of adults between spawning events (kelts). When SAR was greater than approximately 5%, the ratio of true $S_{MSY}$ to a semelparous approximation was only slightly greater than 1 + $e^{-M}$, whereas the ratio quickly increased in a non-linear manner relative to $e^{-M}$ as SAR decreased towards zero (Figure 6b).

### ***Comparison of Capacity and $S_{MSY}$ Estimates with Previous Work***

In order to compare estimates of smolt capacity in this study with those of Gibbons et al. (1985), median estimates of smolt capacity (Table 3) were converted into smolts per 100 $m^2$ using channel surface area values and Beverton-Holt estimates of $S_{MSY}$ were divided by smolt capacity to generate an MSY multiplier ($S_{MSY}$ per smolt at capacity; as reported in Gibbons et al., 1985) at various levels of smolt to adult survival (Table 7). Smolt capacities ranged from 0.29-3.86 smolts per 100 $m^2$ and 0.25-2.56 smolts per 100 $m^2$ for Beverton-Holt and hockey-stick models, respectively, which were very comparable to, but slightly lower overall than, estimates provided by Gibbons et al. (1985), which ranged from 1.83- 3.96 (Table 7).

Our estimates of $S_{MSY}$ per smolt capacity were comparable to those of Gibbons et al. (1985) but differed depending on the model and the marine survival value. Using the Beverton-Holt model, the number of spawners at MSY per 100 $m^2$ of surface area ranged from 0.005-0.006 at 2% smolt to adult survival, to 0.033-0.038 at 15% marine survival, compared with the estimate from Gibbons et al. (1985) of 0.0265 (Table 7).

## 1.E. Discussion

### *Habitat Covariates and Capacity*

Steelhead smolt capacity scaled linearly with stream length and steelhead capacity per unit of stream length was positively associated with PC1, suggesting that both habitat quantity and quality influence capacity. The linear scaling of capacity with stream length likely results because stream length approximates the amount of habitat available for spawning and rearing. Stream length is likely a good measure of habitat availability despite considerable differences in channel size among populations because steelhead juveniles may be confined to margin habitat in large rivers (Hagen et al. 2010), and the results are therefore comparable to previous studies that found associations between habitat quantity and capacity in other species (Chinook; Liermann and 2010, Coho; Barrowman and Myers 2003 and Bradford 2000).

The positive association between smolt capacity and PC1 suggests that capacity per stream kilometer varies predictably with the geomorphological attributes of each watershed. Watershed slope (hillslope gradient) most heavily loaded onto PC1, suggesting that the steelhead smolt capacity per kilometer is higher in watersheds with higher gradient catchments (e.g., hills and mountains) than those flowing from low topographic relief headwaters (e.g., lakes, wetlands, terraces), at least within the range of gradients considered in our study. This may be related to hydrology and sediment transport dynamics because steeper catchments may generate more and larger alluvial sediment and flashier stream flows. Greater alluvial sediment loads could benefit steelhead by creating spawning habitat or promoting productive channel forms (e.g., anabranching, anastomosing channels), while larger substrate sizes such as cobble and even colluvium could benefit steelhead by creating current breaks that facilitate their focal foraging strategy and through the provision of cover (Bjornn 1971). PC1 was also positively associated with the log ratio of mean annual discharge to river length, channel gradient, and channel width, suggesting the after controlling for stream length, rivers with greater discharge, wider channels, and higher gradient had higher smolt capacity. Greater discharge and wider channels may serve as additional surrogates for habitat quantity after stream length, but the positive association with higher gradient suggests that the lowest gradient streams may be less productive for steelhead than for other species such as coho salmon. This is possibly explained by the morphology and feeding behavior of steelhead, which is suited for higher gradient habitats than other stream rearing salmonids and may yield them competitive advantages in these habitats (Bisson et al. 1988).

### *Capacity, Productivity, $S_{MSY}$*

Estimates of capacity based on hockey stick models in this study were somewhat lower than estimates from the Beverton-Holt model. Hockey stick estimates of $S_{MSY}$ were very similar to, but marginally greater than Beverton-Holt estimates of $S_{MSY}$ at 5% marine survival when steelhead were assumed to be semelparous (Table 4). Given the extreme sensitivity of Beverton-Holt estimates $S_{MSY}$ to both smolt-to-adult survival and kelt survival, the hockey-stick estimates of $S_{MSY}$ may be a useful alternative to managers. Not only are hockey stick estimates of $S_{MSY}$ insensitive to smolt and kelt survival but because hockey stick $S_{MSY}$ is equal to $S^*$, use of that reference point ensures freshwater habitat is being adequately saturated, whereas the Beverton-Holt reference point does not. Future refinements to modeled estimates of $S_{MSY}$ using the Beverton-Holt model could be made by incorporating empirical data on adult age composition, harvest removals, and the proportion of adults originating from smolts above traps (where traps are located above the lowermost freshwater rearing habitat). Together these data would enable fitting full spawner to smolt to spawner Integrated Population Models (e.g., Buhle et al. 2018), which would permit direct estimation of population-specific estimates of $S_{MSY}$ via forward projections and calculation of yield under different spawner escapements rather than via approximations (e.g., equations 39-41). Future work could also consider interactions between steelhead and sympatric non-anadromous *O. mykiss* (e.g., rainbow trout) which could affect steelhead populations both by producing additional steelhead smolts, thereby boosting productivity and capacity, as well as by producing resident individuals that compete with rearing steelhead, thereby reducing productivity or capacity (Kendall et al., 2015).

### *Comparison of Capacity and $S_{MSY}$ Estimates with Previous Work*

Smolt capacities in our study were considerably more variable among watersheds, and somewhat lower than estimates provided by Gibbons et al. (1985). Differences between estimates of smolt capacity per unit surface area in two studies may be explained by differences in several modeling aspects of the two studies. The simplest explanation is that Gibbons' work occurred in different watersheds. However, Gibbons et al. (1985) also estimated parr rather than smolt capacity, which should result in higher values because parr must undergo additional overwinter mortality before smolting, and because some parr may spend an additional year in freshwater before smolting. This difference would tend to favor systematically higher estimates of capacity. In contrast, capacities in Gibbons et al. (1985) were based on field observations during years when populations were presumed to be fully seeded, whereas our estimates of capacity (the $\beta$ parameter for Beverton-Holt) represent the asymptotic maximum smolt production given unlimited spawners, which is somewhat greater than what would be expected even with high levels of observable spawner abundance.

Estimates of $S_{MSY}$ per unit of habitat area using the Beverton-Holt model were also more variable in our study than in Gibbons et al. (1985), which we demonstrated was largely a result of sensitivity to smolt and kelt survival. Their results were generally more similar to estimates in our study assuming 15% marine survival, which implies that marine survival in the years they collected data may have been nearer to the highest smolt to adult survival (0.15) than the lowest smolt to adult survival (0.02) evaluated in this study. This is unsurprising because smolt to adult survival has systematically declined throughout coastal Washington since their study was conducted (Kendall et al. 2017). Our hockey-stick model produced a range of spawners at MSY per 100 $m^2$ of 0.013-0.024, which was below that reported by Gibbons et al. (1985), implying that the spawner abundance required to fully seed the habitat ($S*$) is less than that required to maximize yields when marine survival is high. Finally, to facilitate comparison with Gibbons et al. (1985) we reported our estimates using a value of kelt to repeat spawner survival of zero because Gibbons et al. (1985) did not account for iteroparity in estimating $S_{MSY}$; however, Figure 6b provides a visualization of how iteroparity would upwardly adjust $S_{MSY}$ in our study. In general, iteroparity has modest effects when smolt to adult survival is high and kelt survival is low, but failing to account for iteroparity can lead to severely biased estimates of $S_{MSY}$ when kelt survival is higher (e.g., >30%) and smolt survival is lower than 2-3%.

### *Smolt Productivity and Implications for Population Viability*

The productivity of any population is an important parameter for understanding its viability because a population's productivity is the density-independent survival that allows the population to persist at low abundances. Despite its importance, this parameter is often particularly hard to estimate because recruitment in most animal populations is rarely observed at low or near-zero abundances. Therefore, hierarchical models such as those presented herein have particular utility for estimating productivity by sharing information across populations, many of which may independently contain too little information with which to estimate productivity in their abundance time series.

The lowest median estimates of smolt productivity reported in this study were 44 and 81 smolts per spawner for the hockey-stick and Beverton-Holt models, respectively, with hierarchical medians of 56 and 97. These values multiplied by smolt to adult survival must equal at least one for a population to be at replacement, assuming there is no iteroparity. If iteroparity is present, the productivity value that allows replacement at a particular smolt to adult survival is even lower. Because Beverton-Holt productivities are estimated as a population's abundance approaches zero, they may not be particularly informative for understanding viability because minimum viability goals are typically well above extinction. Therefore, using the hockey-stick estimates of productivity which were 50 or greater, one may conclude that 2% marine survival is sufficient for populations to be viable at an abundance that allows full seeding of the habitat (i.e., $S^*$). These results should be of particular interest to managers attempting to establish steelhead viability goals that incorporate smolt to adult survival (e.g., Hard et al. 2013). Perhaps not surprisingly, smolt to adult return (SAR) goals for recovered Columbia River Basin populations have been set at 2-6% for the Columbia Basin (NPCC 2009), consistent with the idea that populations able to fully seed their habitat are likely viable.

### ***Habitat Data Limitations***

Modeled smolt capacity estimates, which were a function of habitat covariates, implicitly assumed that the landscape attributes were measured without error. To obtain comparable habitat attribute data for every stream reach in numerous watersheds, it was necessary to use modeled landscape attributes, meaning that uncertainty and possible bias in capacity estimates related to imperfect habitat attribute data are not accounted for. Since many different models are available to estimate landscape attributes, and their predictions may vary, if researchers wish to apply the landscape attribute-based capacity predictions from this study to new watersheds, care should be taken to use landscape attributes calculated in a manner consistent with those in the NetMap TerrainWorks dataset (Benda 2007) to avoid potential bias.

Some habitat attributes, such as bankfull width and gradient (used in the PCA) were likely particularly susceptible to estimation error. For example, modeled bankfull width assumed that all channels were single-thread, and margin habitat surface area calculated in this report assumed single stream margins on either side of the channel, resulting in bias for multi-thread channels. Furthermore, it is unknown whether fish rearing area responds to bankfull width, or other metrics such as summer wetted width (e.g., Gibbons et al. 1985), which could result in bias since the ratio of bankfull width to wetted width is likely variable. Stream gradients (and other metrics) in TerrainWorks relied upon a 10 m digital elevation model, which was too coarse to precisely measure subtle changes in gradient. This likely particularly impacted predictions of the extent of steelhead occupancy in low gradient areas. Efforts were made to correct the modeled steelhead distribution based on anecdotal knowledge before quantifying habitat attributes for each population. However, application of the steelhead distribution model used in this paper to areas other than those used in this modeling effort would benefit from comparing modeled habitat occupancy with observed steelhead distribution, particularly in low gradient areas, in order to eliminate unused or unusable habitat where the model may falsely predict occurrence.

An additional limitation of our study was that habitat attributes used to model capacity also assumed that the spatial extent of occupied habitat by steelhead in each watershed was known. Since the precise extent of steelhead occupancy within the stream network of each watershed was not known, a logistic regression model was used to predict this the upstream extent of occupancy, and because the logistic regression model was used in the process of developing habitat covariates, its uncertainty was not incorporated into the capacity model.

Finally, caution is warranted in interpreting the relationship between landscape attributes and capacity. We demonstrated that steelhead capacity density was greater in steeper watersheds with higher gradient hillslopes and channel gradients. While this pattern was pronounced in our dataset, the relationship may not have always been as strong. Higher gradient watersheds with steeper hillslopes may have a greater portion of confined channels, and areas where instream habitat is created by boulders, colluvium, and bedrock. Areas with steep hillslopes are generally less affected by land use conversion than flatter areas (Lucero et al. 2011), meaning watersheds with flatter catchments and channels are likely to have experienced greater habitat degradation than steeper watersheds. Compounding this issue, lower gradient reaches may have historically provided more resource subsidies to steelhead from spawning and rearing salmon that have also declined from habitat degradation. The covariance between habitat degradation and landscape features (e.g., Lucero et al. 2011) suggests that lower gradient habitats may have once been more productive, and although apparently less productive for steelhead now, may nonetheless offer substantial restoration potential to increase steelhead population capacity.

### ***Effects of Hatchery Spawners***

One potential source of bias and uncertainty in our models was the effect of hatchery-origin spawners on productivity and capacity of wild populations. The number of hatchery origin spawners was highly uncertain for some populations, and precise estimates of hatchery spawner relative reproductive success were unavailable, and these two factors therefore had the potential to affect the accuracy of estimates of wild spawner equivalents used in this analysis. However, the magnitude of this effect is naturally scaled to the proportional contribution of hatchery spawners to total wild spawner abundance (i.e., when hatchery spawners comprised a small proportion of the wild spawner equivalents, the magnitude of any potential effect is small). Additionally, because high hatchery spawner proportions typically resulted in high wild spawner equivalent abundances for a particular population, the effect of potential bias in the estimation of population parameters was inherently reduced. Stock-recruit data points at high spawner abundances are more informative to the capacity parameter, which is mostly a function of smolt abundance, as opposed to stock-recruit data points at low spawner abundances, which heavily inform the productivity parameter, which depends on both spawner and smolt abundance. As a result, it is unlikely that the adjustments to the spawner data to accommodate hatchery spawners had a strong influence on reference points in this analysis.

### *Conclusions and Future Work*

Hierarchical Bayesian hockey-stick and Beverton-Holt models were parameterized to incorporate habitat covariates enabled estimation of population reference points for 14 coastal Washington steelhead populations. This approach also enabled the development of an empirical predictive model that allows for estimation of reference points for unmonitored populations based on their habitat attributes and the density independent smolt to adult and kelt survival at the time. The results suggest that watershed characteristics were strongly associated with steelhead smolt capacity. However, predictive distributions for capacity using these variables alone contained considerable uncertainty. This result highlights the need to explore additional covariates and collect more, and longer, high quality smolt and adult abundance datasets, which should help to understand variability in capacity among populations resulting from factors other than habitat quantity and the limited watershed attributes considered in the PCA. Before such data are available, approaches such as the one used here will enable more robust estimates of biological reference points, particularly in locations where data are limited.

## 1.F. References

Araki, H., B. A. Berejikian, M. J. Ford, and M. S. Blouin. 2008. Fitness of hatchery-reared salmonids in the wild. Evolutionary Applications, 1(2): 342–355.

Barrowman, N. J., R. A. Myers, R. Hilborn, D. G. Kehler, and C. A. Field. 2003. The variability among populations of coho salmon in the maximum reproductive rate and depensation. Ecological Applications 13(3): 784-793.

Barrowman, N. J., R. A. Myers. 2000. Still more spawner-recruitment curves: the hockey stick and its generalizations. Can. J. Fish. Aquat. Sci. 57: 665–676.

Beverton, R. J., and S. J. Holt. 1957. On the dynamics of exploited fish populations. MAFF Fisheries Investigations Series II 19:533.

Beecher, H. A., T. H. Johnson, and J. P. Carleton. 1993. Predicting microdistributions of steelhead (Oncorhynchus mykiss) parr from depth and velocity preference criteria: test of an assumption of the instream flow incremental methodology. Can. J. Fish. Aquat. Sci., 50: 2380-2387.

Benda, L., D. J. Miller, K. Andras, P. Bigelow, G. Reeves, and D. Michael. 2007. NetMap: A new tool in support of watershed science and resource management. Forest Science 52:206-219.

Bisson, P. A., K. Sullivan, and J. L. Nielsen. 1988. Channel hydraulics, habitat use, and body form of juvenile coho salmon, steelhead, and cutthroat trout in streams. Transactions of the American Fisheries Society, 118(3), 467–483.

Bjornn, T. C. 1971. Trout and salmon movements in two Idaho streams as related to temperature, food, stream flow, cover, and population density. Transactions of the American Fisheries Society, 100(3): 423–438.

Bonner, S., and C. J. Schwarz. 2011. Smoothing Population Size Estimates for Time-Stratified Mark–Recapture Experiments Using Bayesian P-Splines. Biometrics, 67(4) 1498–1507. DOI: 10.1111/j.1541-0420.2011.01599.x

Bradford, M. J., G. C. Taylor, and J. A. Allen. 1997. Empirical review of coho salmon smolt abundance and the prediction of smolt production at the regional level. Trans. Amer. Fish. Soc. 126(1): 49-64.

Bradford, M. J. 1999. Temporal and spatial trends in the abundance of coho salmon smolts from Western North America. Transactions of the American Fisheries Society, 128:5, 840-846.

Bradford, M.J., R. A. Myers, and J. R. Irvine. 2000. Reference points for coho salmon harvest rates and escapement goals based on freshwater production. *Canadian Journal of Fisheries and Aquatic Sciences* 57: 677-686.

Brooks, E. N., J. E. Powers, and E. Corte´s. 2010. Analytical reference points for age-structured models: application to data-poor fisheries. ICES Journal of Marine Science, 67: 165–175.

Buhle, E. R., M. D. Scheuerell, T. D. Cooney, M. J. Ford, R. W. Zabel, and J. T. Thorson. 2018. Using Integrated Population Models to Evaluate Fishery and Environmental Impacts on Pacific Salmon Viability. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-NWFSC-140. DOI: 10.7289/V5/TM-NWFSC-14.

Burnett, K. M., G. H. Reeves, D. J. Miller, S. Clarke, K. Vance-Borland, and K. Christiansen. 2007. Distribution of salmon-habitat potential relative to landscape characteristics and implications for conservation. Ecological Applications 17(1): 66-80.

Cramer, S. P., and N. K. Ackerman. 2009. Linking stream carrying capacity for salmonids to habitat features. American Fisheries Society Symposium 71: 225-254.

Darroch, J. N. 1961. The two-sampled capture–recapture census when tagging and sampling are stratified. Biometrika 48:241–260.

Gallagher, S. P., P. K. J. Hahn, and D. H. Johnson. 2007. Redd Counts. Pages 197-234 in D. H. Johnson, B. M. Shrier, J. S. O'Neal, J. A. Knutzen, X. Augerot, T. A. O-Neil, and T. N. Pearsons, editors. Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations. American Fisheries Society, Bethesda, Maryland.

Gayeski, N., G. Pess, and T. Beechie. 2016. A life-table model estimation of the parr capacity of a late 19th century Puget Sound steelhead population. Facets 1: 83–104.

Geweke, J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics,* edited by. J. M. Bernado, J. O. Berger, A. P. Dawid and A. F. M. Smith. Clarendon Press, Oxford, UK.

Gibbons, R. G., P. K. J. Hahn, and T. H. Johnson. 1985. Methodology for determining MSH steelhead spawning escapement requirements. Washington State Game Department, Fisheries Management Division Report 85-11, Olympia, WA, 39 p.

Hagen, J., S. Decker, J. Korman, and R. G. Bison. 2010. Effectiveness of night snorkeling for estimating steelhead parr abundance in a large river basin. North American Journal of Fisheries Management 30, 1303-1314.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall.

J. Hall, P. Roni, T. Bennett, J. McMillan, K. Hanson, R. Moses, M. McHenry, G. Pess, and W. Ehinger. 2016. Life history diversity of steelhead in two coastal Washington watersheds. Transactions of the American Fisheries Society 145(5): 990-1005.

Hard, J. H., J. M. Myers, E. J. Connor, R.A. Hayman, R. G. Kope, G. Lucchetti, A. R. Marshall, G. R. Pess, and B. E. Thompson. 2013. Viability criteria for Puget Sound Steelhead. Puget Sound Steelhead Technical Recovery Team (PSSTRT). Final Review Draft. 373 p.

Hilborn, R., and C. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics, and uncertainty. Chapman and Hall, New York. 570 p.

Johnson, N. T., E.A. Parkinson, A. F. Tautz, and B. R. Ward. 2000. Biological reference points for the conservation and management of steelhead, *Oncorhynchus mykiss*. Canadian Stock Assessment Secretariat, 2000/126, 96 pp.

Jonsson, N., B. Jonsson, and L. P. Hansen. 1998. The relative role of density-dependent and density-independent survival in the life cycle of Atlantic salmon *Salmo salar*. Journal of Animal Ecology, 67: 751–762. doi:10.1046/j.1365-2656.1998.00237.x

Keeley, E. R. 2000. An experimental analysis of territory size in juvenile steelhead trout. Animal Behaviour 59, 477–490.

Kendall, N. W., J. R. McMillan, M. R. Sloat, T. W. Buehrens, T. P. Quinn, G. R. Pess, K. V. Kuzishchin, M. M. McClure, R. W. Zabel. 2015. Anadromy and residency in steelhead and rainbow trout (Oncorhynchus mykiss): a review of the processes and patterns. Canadian Journal of Fisheries and Aquatic Sciences 72: 319-342.

Kendall, N. W., G. W. Marston, and M. M. Klungle. 2017. Declining patterns of Pacific Northwest steelhead trout (Oncorhynchus mykiss) adult abundance and smolt survival in the ocean. Canadian Journal of Fisheries and Aquatic Sciences 74(8):1275-1290.

Kostow, K. E., A. R. Marshall, and S. R. Phelps. 2003. Naturally spawning hatchery steelhead contribute to smolt production but experience low reproductive success. Transactions of the American Fisheries Society 132: 780–790.

Liermann, M. C., R. Sharma, and C. K. Parken. 2010. Using accessible watershed size to predict management parameters for Chinook salmon, *Oncorhynchus tshawytscha*, populations with little or no spawner-recruit data: a Bayesian hierarchical modelling approach. Fisheries Management and Ecology 17: 40-51.

Lucero Y., E. A. Steel, K. M. Burnett, and K. Christiansen. 2011. Untangling human development and natural gradients: implications of underlying correlation structure for linking landscapes and riverine ecosystems. River Syst,19(3): 207-224.

Ludwig, D. and C. J. Walters. 1981. Measurement errors and uncertainty iun parameter estimates for stock and recruitment. Can. J. Fish. Aquat. Sci.38: 711:720.

McCormick, J. L. and J. R. Ruzycki. 2016. A Simulation Assessment of

Model-Assisted Estimation of Steelhead Redd Abundance, North American Journal of Fisheries Management, 36:4, 913-925, DOI: 10.1080/02755947.2016.1173140

Michielsens, C. G. J., and M. K. McAllister. 2004. A Bayesian hierarchical analysis of stock recruit data: quantifying structural and parameter uncertainties. Can. J. Fish. Aquat. Sci. 61: 1032–1047.

Milner, N. J., J. M. Elliott, J. D. Armstrong, R. Gardiner, J. S. Welton, and M. Ladle. 2003. The natural control of salmon and trout populations in streams. Fisheries Research 62:111–125.

Myers, R.A., A. A. Rosenberg, P. M. Mace, N. J. Barrowman, and V. R. Restrepo. 1994. In search of thresholds for recruitment overfishing. ICES J. Mar. Sci. 51: 191–205.

Myers, R. A., G. Mertz, and J. Bridson. 1997. Spatial scales of interannual recruitment variations of marine, anadromous, and freshwater fish. Canadian Journal of Fisheries and Aquatic Sciences 54:1400–1407.

NPCC (Northwest Power and Conservation Council). 2009. Columbia River Basin Fish and 14 Wildlife Program. Council Document 2009-02. http://www.nwcouncil.org/library/2009/2009-02.pdf

Peterman, R.M., B. J. Pyper, M. F. Lapointe, M. D. Adkison and C. J. Walters. 1998. Patterns of covariation in survival rates of British Columbian and Alaskan sockeye salmon (*Oncorhynchus nerka*) stocks. Canadian Journal of Fisheries and Aquatic Sciences 55: 2503-2517.

Plummer, M. 2003. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling, Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), March 20–22, Vienna, Austria.

Quinn, T.P. 2005. The Behavior and Ecology of Pacific Salmon and Trout. University of Washington Press.

Rivot, E., E. Prévost, E. Parent, and J. L. Baglinière. 2004. A Bayesian state-space modelling framework for fitting a salmon stage-structured population dynamic model to multiple time series of field data. Ecological Modelling 179(4): 463-485.

Schnute J.T., and A. R. Kronlund. 2002. Estimating salmon stock-recruit relationships from catch and escapement data. Canadian Journal of Fisheries and Aquatic Sciences 59, 433–449.

Sharma, R. and R. Hilborn. 2001. Empirical relationships between watershed characteristics and coho salmon (*Oncorhynchus kisutch*) smolt abundance in 14 western Washington streams. *Canadian Journal of Fisheries and Aquatic Sciences* 58: 1453-1463.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde. 2002. Bayesian measures of model complexity and fit (with discussion). Journal of the Royal Statistical Society B64:582-639.

Strahler, A. N. 1952. Hypsometric (area-altitude) analysis of erosional topology, Geological

Society of America Bulletin, 63 (11): 1117–1142.

Su, Y-S., and M. Yajima. 2015. R2jags: using R to run JAGS. R package, version 0.5-7. http://CRAN.R-project.org/package=R2jags.

Tautz, A. F., B. R. Ward, and R. A. Ptolemy. Steelhead trout productivity and stream carrying capacity for rivers of the Skeena drainage. PSARC Working Paper S92-6 and 8.

Volkhardt, G.C., Johnson, S.L., Miller, B.A., Nickelson, T.E. & Seiler, D.E. 2007. Rotary screw traps and inclined plane screen traps. In: Johnson, D.H., Shrier, B.M., O'Neal, J.S., Knutzen, J.A., Augerot, X., O-Neil, T.A. & Pearsons, T.N., eds. *Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations*. Bethesda, Maryland: American Fisheries Society, pp. 235-266.

Ward, B. R. and P. A. Slaney. 1993. Egg-to-smolt survival and fry-to-smolt density dependence of Keogh River Steelhead Trout. In R. J. Gibson and R. E. Cutting [ed.] Production of juvenile Atlantic salmon, *Salmo Salar*, in natural waters. Can. Spec. Publ. Fish. Aquati. Sci. 118.

Ward, B. R. 2000. Declivity in steelhead (Oncorhynchus mykiss) recruitment at the Keogh River over the past decade. Can. J. Fish. Aquat. Sci. 57(2): 298-306.

Zimmerman, M. S., J. R. Irvine, M. O’Neill, J. H. Anderson, C. M. Greene, J. Weinheimer, M. Trudel, and K. Rawson. 2015. Spatial and temporal patterns in smolt survival of wild and hatchery coho salmon in the Salish Sea. Marine and Coastal Fisheries **7**(1): 116-134.

### 1.G. Acknowledgements

Countless staff labored to collect smolt and adult abundance datasets and juvenile scales used in our analysis. Entire careers were devoted to these monitoring datasets, for which the idiom “the sum is greater than the parts” is truer than ever. We are deeply indebted to these long-term monitoring programs and the data streams they generate. Specific thanks are due to the following WDFW staff who collected smolt and adult data: Mark Downen, Clayton Kinsel, Patrick “Charlie” Cochran, Jamie Lamperth, Bryce Glaser, and Steve Gray. Additionally, Steven Vanderploeg at WDFW helped with creating maps and managing spatial datasets. Finally, we thank Mara Zimmerman of the Coast Salmon Partnership and formerly of WDFW, Lukas DeFilipo (NOAA Fisheries) and Eric Buhle (Mt. Hood Environmental) for valuable conversations that helped guide this work.

### 1.H. Tables

Table 1. Wild steelhead spawner and smolt datasets used to estimate population reference points. Run types include summer (S) and winter (W) run steelhead. Distinct Population Segment (DPS) describes evolutionary unit managed for species recovery under the US Endangered Species Act. The methods of abundance estimation and the number of years of available data are provided for each population; MR = mark-recapture, C = census (weirs/dams), and R = redd counts. The column "Effective hatchery spawners" lists the range among years of contributions of hatchery spawners to the total "wild spawner equivalent" abundance for each population (see methods).

| Population | Run Type | DPS | Spawner Method | Yrs. | Effective Hatchery Spawners | Smolt Method | Yrs. |
|---|---|---|---|---:|---:|---|---:|
| Abernathy | W | Lower Columbia | R | 17 | 7-49% | MR | 14 |
| Big Beef | W | Puget Sound | R & C | 5 | 0% | C | 7 |
| Bingham | W | Southwest Washington | C | 16 | 0% | C | 13 |
| Cedar | W | Lower Columbia | MR | 18 | 7-51% | MR | 15 |
| Coweeman | W | Lower Columbia | R | 13 | 0-15% | MR | 8 |
| EF Lewis | W & S | Lower Columbia | R & MR | 4 | 27-61% | MR | 1 |
| Germany | W | Southwest Washington | R | 18 | 7-21% | MR | 15 |
| Grays | W | Southwest Washington | R | 10 | 1-32% | MR | 6 |
| Kalama | W & S | Lower Columbia | MR & C | 38 | 1-65% | MR | 23 |
| Mill | W | Southwest Washington | R | 16 | 0% | MR | 15 |
| NF Toutle | W | Lower Columbia | C | 2 | 0% | MR | 1 |
| Snow Creek | W | Puget Sound | C | 37 | 0% | C | 36 |
| Trout | S | Lower Columbia | C & MR | 20 | 0% | MR | 21 |
| Wind | S | Lower Columbia | MR | 22 | 0-33% | MR | 19 |

Table 2. Habitat attributes for stream reaches used by coastal steelhead populations above smolt trapping locations. Normalized and standardized combinations (see methods) of these variables were used in a PCA describing differences in habitat among watersheds. PC1 accounted for 50% of the variation in the habitat dataset (see results).

| Population | Occupied Stream Length (km) | Mean Annual Discharge ($m^3 \cdot s^{-1}$) | Channel Gradient (% slope) | Cumulative Hillslope Gradient (% slope) | Channel Width (m) | PC1 loading |
|---|---|---|---|---|---|---|
| Abernathy Creek | 28.6 | 4.4 | 1.96 | 29 | 7.8 | 0.14 |
| Big Beef Creek | 8.4 | 1.0 | 1.12 | 11 | 15.0 | -1.56 |
| Bingham Creek | 36.7 | 5.0 | 0.45 | 9 | 6.9 | -3.18 |
| Cedar Creek | 45.3 | 6.8 | 0.73 | 19 | 8.6 | -1.41 |
| Coweeman River | 73.2 | 17.1 | 1.19 | 29 | 13.2 | 0.51 |
| EF Lewis River | 96.0 | 31.3 | 1.18 | 26 | 16.2 | 0.85 |
| Germany Creek | 17.8 | 3.4 | 1.33 | 28 | 10.5 | 0.20 |
| Grays River | 66.9 | 20.5 | 1.56 | 39 | 11.0 | 1.45 |
| Kalama River | 97.4 | 42.4 | 1.35 | 35 | 15.4 | 1.73 |
| Mill Creek | 29.0 | 4.3 | 2.19 | 18 | 7.3 | -0.59 |
| NF Toutle River | 67.1 | 28.6 | 1.44 | 32 | 16.0 | 1.64 |
| Snow Creek | 11.2 | 1.1 | 1.75 | 23 | 6.1 | -0.99 |

Table 3. Posterior median smolt capacity per kilometer based on the Beverton-Holt and hockey stick models. Values are posterior medians and 80% credible intervals. Results are reported for each population both as models fits as well as based on habitat data only using the posterior predictive distribution with PC1 scores for each population. Statistics are also reported for the posterior predictive distribution for PC1 score quantiles of 0.10, 0.50, and 0.90). This represents what the model would predict for each population without an estimate of the population-specific random effect, enabled by inclusion of its fish data in the model.

|  | BH Model Fit |  | BH Posterior Predictive Distribution Without Fish Data |  | HS Model Fit |  | HS Posterior Predictive Distribution Without Fish Data |  |
|---|---|---|---|---|---|---|---|---|
| Population | median | 80% CI | median | 80% CI | median | 80% CI | median | 80% CI |
| Abernathy Creek | 125 | (100-184) | 198 | (102-378) | 99 | (84-129) | 148 | (84-265) |
| Big Beef Creek | 124 | (94-167) | 79 | (40-155) | 99 | (80-123) | 63 | (35-116) |
| Bingham Creek | 26 | (22-31) | 33 | (15-72) | 22 | (19-25) | 28 | (14-56) |
| Cedar Creek | 55 | (47-64) | 86 | (43-168) | 48 | (43-54) | 68 | (38-125) |
| Coweeman River | 337 | (258-480) | 241 | (124-463) | 232 | (199-283) | 179 | (101-321) |
| EF Lewis River | 333 | (202-588) | 289 | (147-561) | 240 | (160-387) | 212 | (119-384) |
| Germany Creek | 376 | (299-498) | 203 | (105-388) | 275 | (241-314) | 152 | (86-273) |
| Grays River | 306 | (224-460) | 396 | (198-788) | 223 | (180-286) | 285 | (157-531) |
| Kalama River | 321 | (280-379) | 459 | (228-931) | 273 | (248-300) | 328 | (178-618) |
| Mill Creek | 97 | (68-169) | 133 | (69-255) | 73 | (50-157) | 103 | (58-184) |
| NF Toutle River | 427 | (201-873) | 440 | (219-884) | 315 | (172-577) | 314 | (171-590) |
| Snow Creek | 137 | (115-165) | 107 | (55-208) | 109 | (94-127) | 84 | (47-151) |
| Trout Creek | 219 | (177-281) | 189 | (98-361) | 137 | (123-154) | 142 | (81-255) |
| Wind River | 328 | (265-441) | 338 | (170-663) | 214 | (194-241) | 245 | (137-449) |
| Pred. Dist. PC1 Q 0.10 | NA | NA | 92 | (47-179) | NA | NA | 73 | (40-132) |
| Pred. Dist. PC1 Q 0.50 | NA | NA | 183 | (95-350) | NA | NA | 138 | (78-247) |
| Pred. Dist. PC1 Q 0.90 | NA | NA | 363 | (183-715) | NA | NA | 263 | (146-485) |

Table 4. Posterior median estimates of productivity ($\alpha$), the maximum per-capita smolt recruitment, for 14 coastal steelhead populations based on Beverton-Holt and hockey stick models. Data are medians and 80% credible intervals of the posterior distribution. The “unknown” population is the posterior prediction for a random population whose parameters are unknown.

| | Beverton-Holt | | Hockey Stick | |
|---|---|---|---|---|
| Population | median | 80% CI | median | 80% CI |
| NF Toutle River | 81 | (47-119) | 44 | (35-62) |
| Mill Creek | 84 | (54-127) | 51 | (40-64) |
| Abernathy Creek | 85 | (36-136) | 52 | (35-73) |
| Wind River | 90 | (64-130) | 54 | (19-86) |
| Grays River | 93 | (53-156) | 54 | (41-70) |
| Big Beef Creek | 97 | (66-157) | 55 | (33-89) |
| Cedar Creek | 97 | (56-171) | 56 | (36-90) |
| Bingham Creek | 98 | (59-170) | 56 | (35-90) |
| EF Lewis River | 100 | (63-177) | 56 | (36-90) |
| Coweeman River | 104 | (70-197) | 57 | (38-92) |
| Kalama River | 105 | (68-203) | 57 | (40-82) |
| Snow Creek | 105 | (76-166) | 58 | (39-95) |
| Germany Creek | 112 | (76-232) | 59 | (42-95) |
| Trout Creek | 112 | (86-156) | 75 | (60-94) |
| Unknown Population | 97 | (57-172) | 56 | (35-91) |

Table 5. Posterior predictive distributions of the spawner density that produces the maximum sustainable yield for the hockey-stick (HS) and Beverton-Holt model (BH) models when repeat spawning does not occur (e.g., steelhead are semelparous). Values are reported for three levels of smolt to adult survival (SAR = 2%, 5%, and 15%) and for the 10th, 50th, and 90th quantiles of habitat productivity as measured by PC1. See Figure 6 for approximate bias corrections to $S_{MSY}$ under varying levels of iteroparity.

| | | | Spawner/km | |
|---|---|---|---|---|
| PC1 Quantile (Value) | Model | SAR | Median | (80% CI) |
| 0.10 (-1.28) | BH | 0.02 | 0.33 | (0.08-0.69) |
| 0.10 (-1.28) | BH | 0.05 | 1.08 | (0.52-2.14) |
| 0.10 (-1.28) | BH | 0.15 | 2.61 | (1.27-5.26) |
| 0.50 (0) | BH | 0.02 | 0.66 | (0.17-1.36) |
| 0.50 (0) | BH | 0.05 | 2.15 | (1.06-4.16) |
| 0.50 (0) | BH | 0.15 | 5.19 | (2.55-10.33) |
| 0.90 (1.28) | BH | 0.02 | 1.31 | (0.32-2.78) |
| 0.90 (1.28) | BH | 0.05 | 4.27 | (2.06-8.57) |
| 0.90 (1.28) | BH | 0.15 | 10.33 | (4.95-21.16) |
| 0.10 (-1.28) | HS | NA | 1.29 | (0.58-2.90) |
| 0.50 (0) | HS | NA | 2.45 | (1.12-5.41) |
| 0.90 (1.28) | HS | NA | 4.65 | (2.09-10.52) |

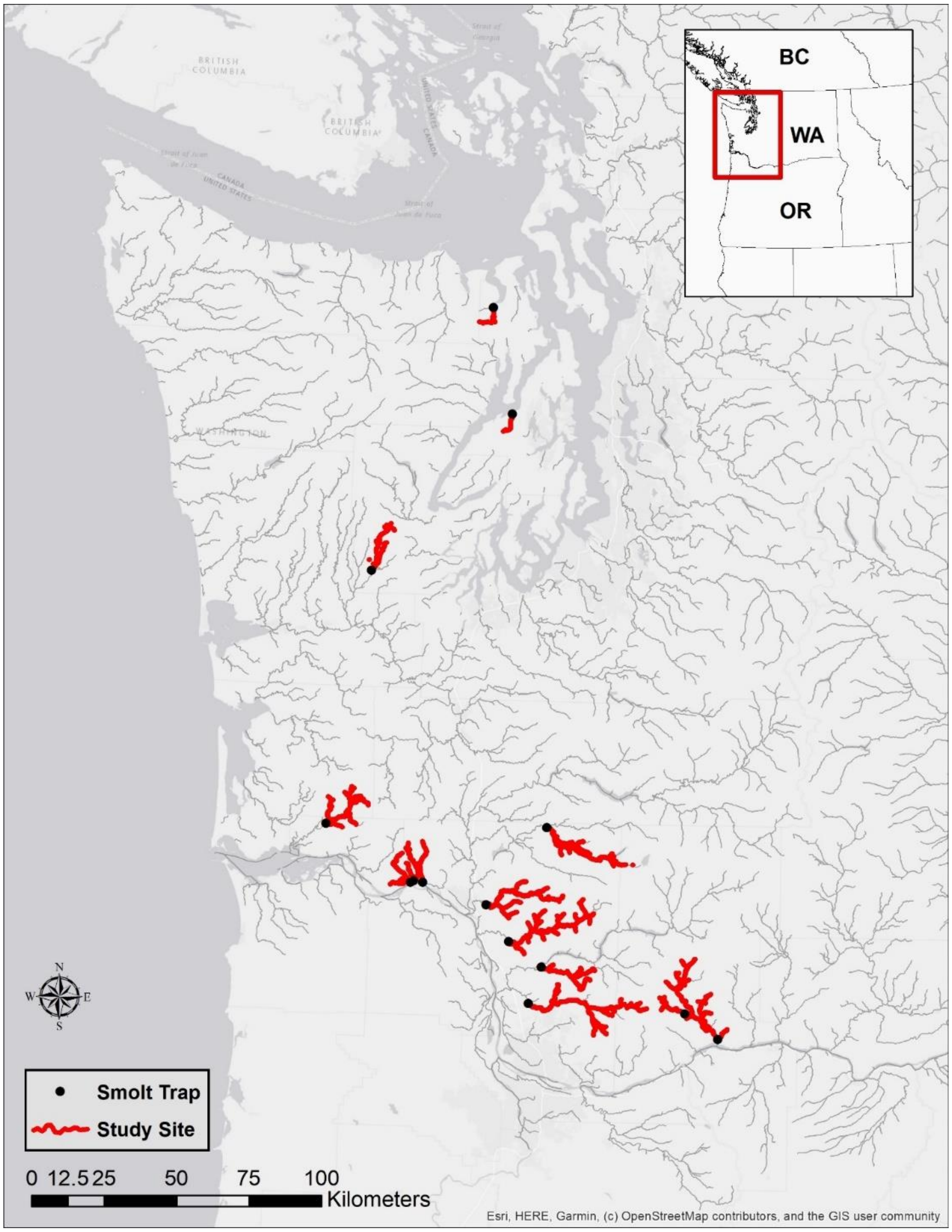
Table 6. Median estimates of smolt capacity per 100 $m^2$ of surface area and $S_{MSY}$ per smolt at capacity at various levels of smolt to adult survival.

| Population | Surface Area (100 $m^2$) | Capacity per 100 $m^2$ (smolts *100 $m^{-2}$) | | MSY Multiplier (Semelparous $S_{MSY}$ per smolt capacity) | | | |
|---|---|---|---|---|---|---|---|
| | | BH | HS | HS | BH SAR 0.02 | BH SAR 0.05 | BH SAR 0.15 |
| Abernathy Creek | 2,183 | 1.64 | 1.30 | 0.018 | 0.005 | 0.016 | 0.036 |
| Big Beef Creek | 1,303 | 0.80 | 0.64 | 0.019 | 0.006 | 0.016 | 0.035 |
| Bingham Creek | 3,283 | 0.29 | 0.25 | 0.018 | 0.006 | 0.015 | 0.034 |
| Cedar Creek | 4,260 | 0.58 | 0.51 | 0.018 | 0.006 | 0.015 | 0.035 |
| Coweeman River | 10,854 | 2.27 | 1.56 | 0.017 | 0.006 | 0.015 | 0.034 |
| EF Lewis River | 19,501 | 1.64 | 1.18 | 0.018 | 0.006 | 0.015 | 0.034 |
| Germany Creek | 1,913 | 3.50 | 2.56 | 0.017 | 0.006 | 0.015 | 0.033 |
| Grays River | 8,345 | 2.45 | 1.79 | 0.018 | 0.006 | 0.016 | 0.035 |
| Kalama River | 19,674 | 1.59 | 1.35 | 0.018 | 0.006 | 0.015 | 0.034 |
| Mill Creek | 2,089 | 1.35 | 1.01 | 0.024 | 0.006 | 0.016 | 0.037 |
| NF Toutle River | 17,390 | 1.65 | 1.22 | 0.019 | 0.005 | 0.016 | 0.038 |
| Snow Creek | 698 | 2.20 | 1.75 | 0.017 | 0.006 | 0.016 | 0.034 |
| Trout Creek | 1,212 | 3.38 | 2.11 | 0.013 | 0.006 | 0.015 | 0.033 |
| Wind River | 9,994 | 3.86 | 2.52 | 0.019 | 0.006 | 0.016 | 0.036 |

Table 7. Estimated steelhead parr production potential, habitat surface area, and parr density at assumed capacity from Gibbons et al. (1985). Nisqually River parr capacity and $S_{MSY}$ were reported but no surface area was listed.

| Population | Gibbons Parr Capacity | Surface Area (100 m$^2$) | Parr*100 m$^{-2}$ | $S_{MSY}$ | MSY Multiplier (Semelparous $S_{MSY}$ per smolt capacity) |
|---|---|---|---|---|---|
| Quillayute | 223,084 | 80,879 | 2.76 | 5912 | 0.0265 |
| Queets | 157,239 | 63,363 | 2.48 | 4167 | 0.0265 |
| Skagit | 387,672 | 211,299 | 1.83 | 10273 | 0.0265 |
| Green | 76,142 | 27,871 | 2.73 | 2018 | 0.0265 |
| Nisqually | 65,813 | Not reported | Not reported | 1744 | 0.0265 |
| Kalama | 74,924 | 18,933 | 3.96 | 1985 | 0.0265 |

## 1.I. Figures



Esri, HERE, Garmin, (c) OpenStreetMap contributors, and the GIS user community

Figure 1. Map of Western Washington showing the location of the fourteen steelhead smolt traps above which steelhead spawner and smolt abundance estimates used in this study were made. Stream lines that are bolded and colored red show the reaches estimated to be occupied by steelhead and for which habitat attributes were calculated for use in estimating watershed effects on capacity.

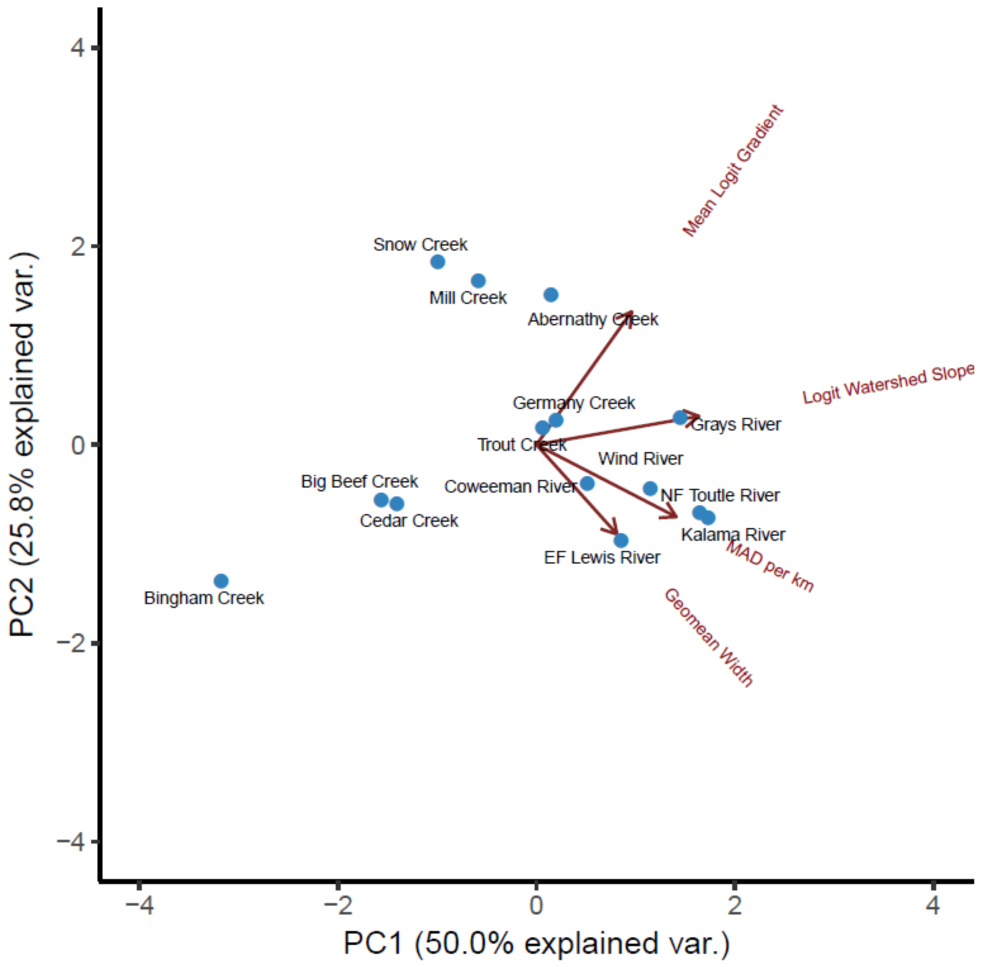


Figure 2. Scatterplot of PC1 and PC2 values for each of the 14 coastal watersheds. Arrows show loadings of four habitat attributes hypothesized to be associated with steelhead smolt capacity.

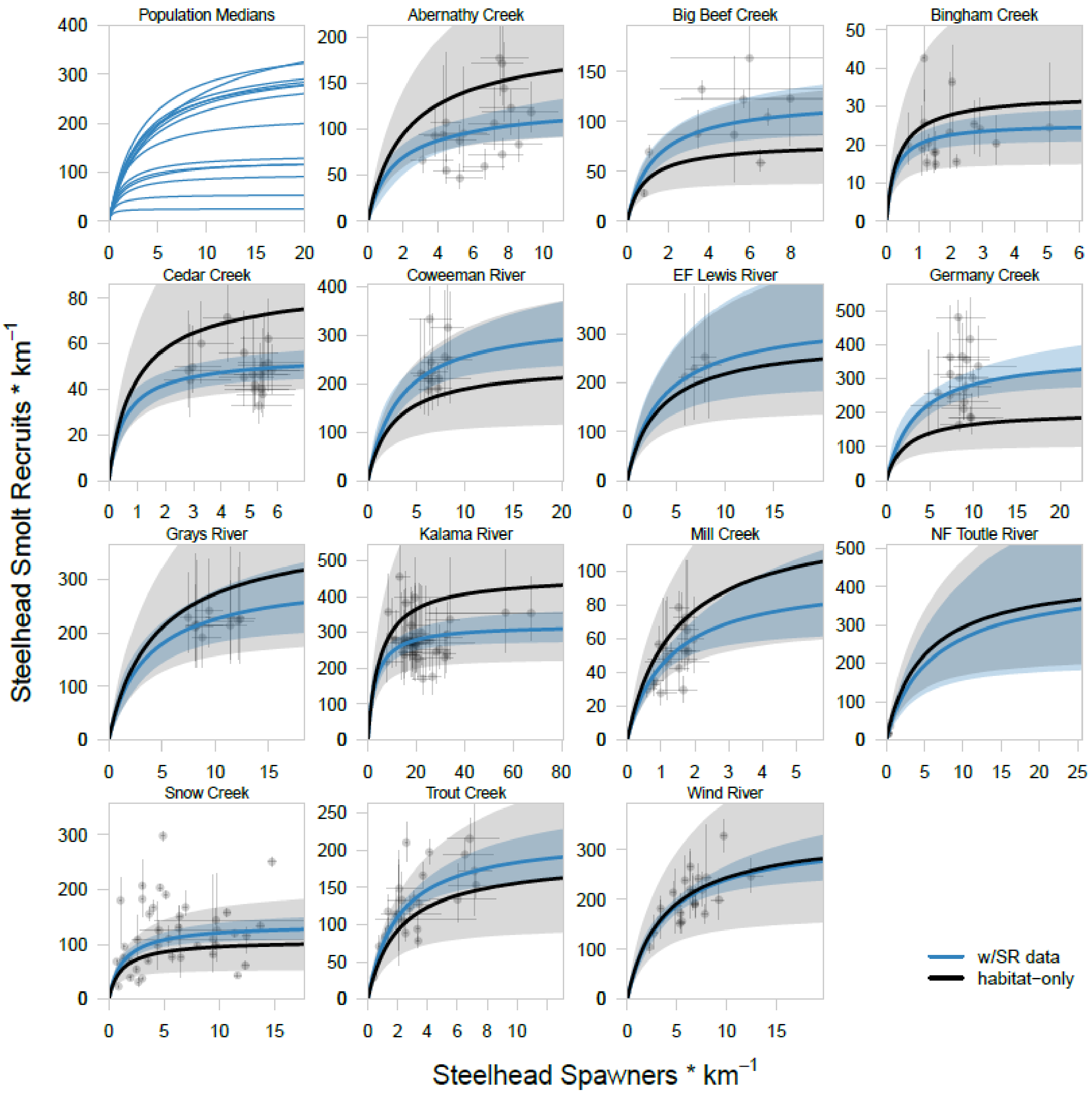


Figure 3. Estimated wild steelhead smolt recruits plotted against estimated brood year spawner abundance for 14 coastal populations in Washington State. Points are medians and grey lines show 90% CI on both axes for the estimates of each spawner-recruit pair. Blue lines are the posterior median estimate of recruitment for the Beverton-Holt model of the stock-recruitment relationship for each population. Black lines are the estimated recruitment based on the posterior predictive distribution (drawing random effects estimated in equation 11 from the hyper-distribution with mean 0 for each population). Shading indicates 80% credible intervals. Hockey-stick results are shown in Figure S1).

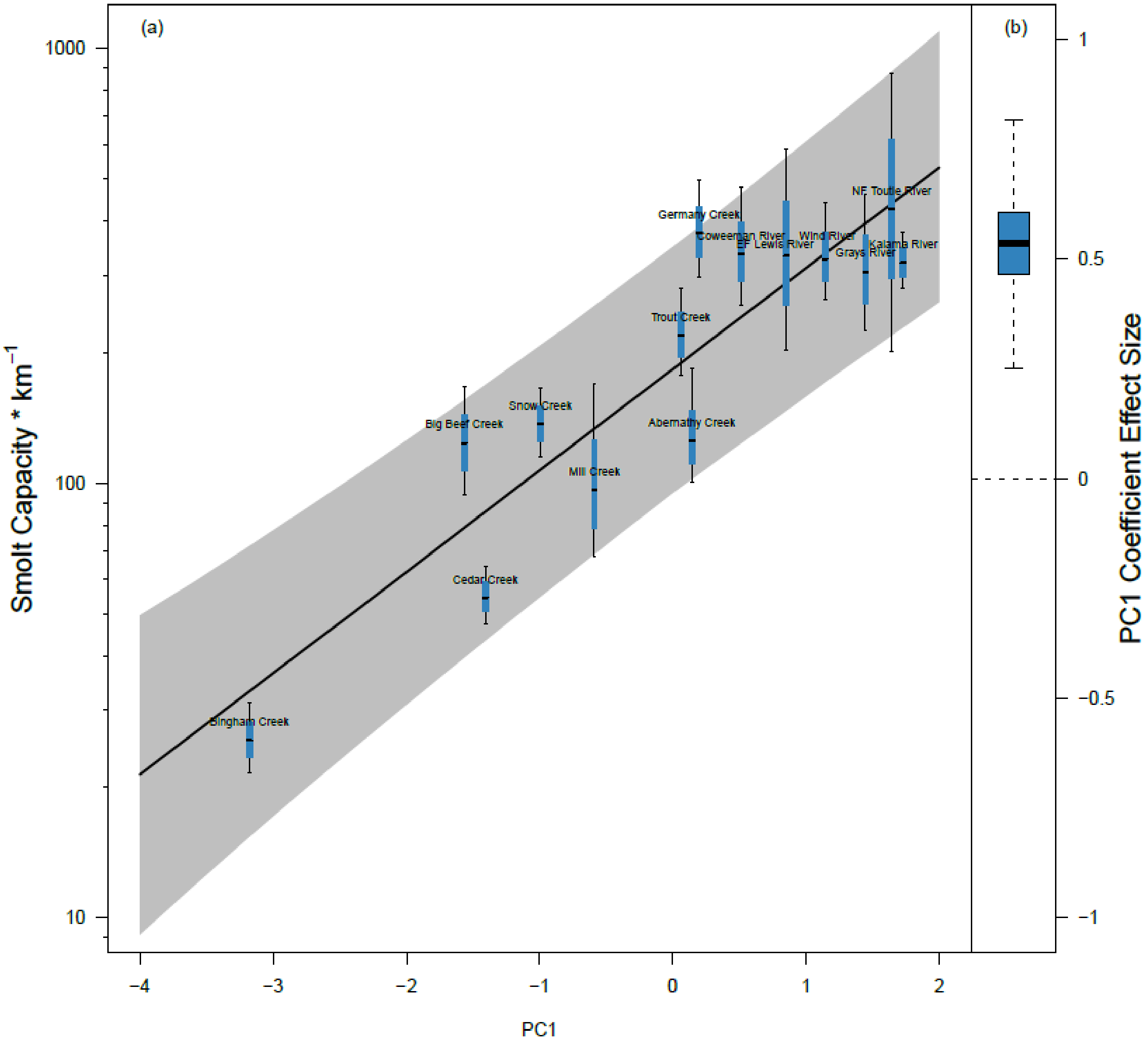


Figure 4. The relationship between estimated smolt capacity (*αS** or *K*) per kilometer of habitat and the habitat PC1 for the Beverton-Holt model for 14 coastal steelhead populations (a), and the posterior distribution for the coefficient determining the effect of habitat PC1 on capacity (b). Blue lines and whiskers are medians and 80% credible intervals of the posterior distribution for each population. The black line and grey polygon show the median and 80% credible interval for the posterior predictive distribution for an unknown population. Note log scales on y-axis of (panel a). Hockey-stick results are shown in Supplementary Figure 2).

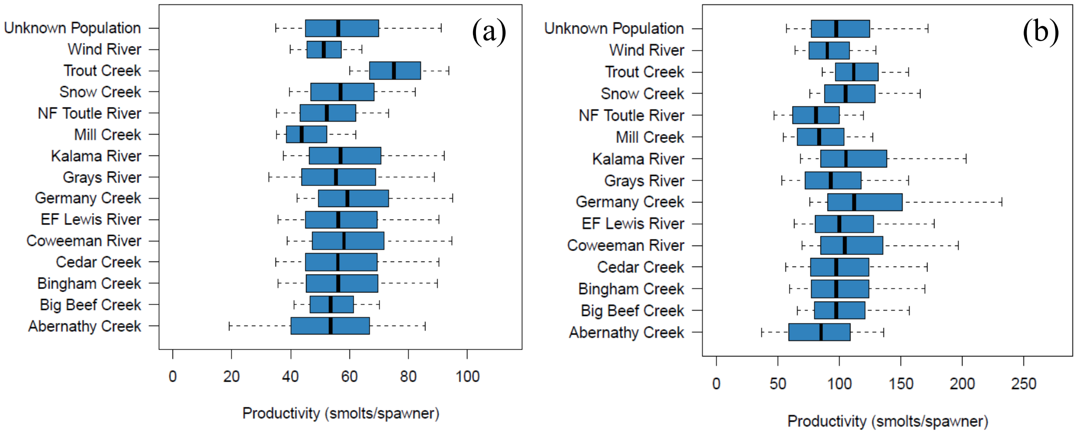


Figure 5. Estimated asymptotic per-capita recruitment ($\alpha$) for 14 coastal steelhead populations based on the Beverton-Holt stock recruit model (a) and hockey stick model (b). Values are medians and boxes and whiskers are 50% and 80% credible intervals of the posterior distribution, respectively.

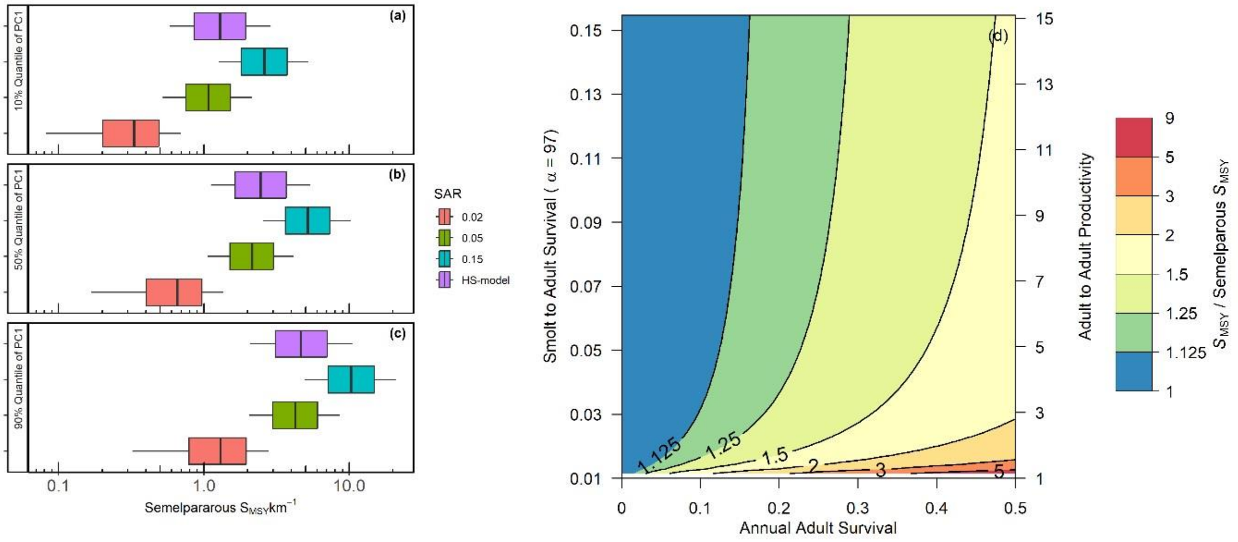


Figure 6. Panel (a): Posterior predictive distributions of the spawner density needed to achieve produce the maximum sustainable yield ($S_{MSY}$) for the hockey stick and Beverton-Holt model when repeat spawning does not occur (i.e., steelhead are semelparous). Boxplots show medians and boxes and whiskers are 50% and 80% credible intervals of the posterior distribution, respectively. Values are reported for three levels of smolt to adult survival (2%, 5%, and 15%) and for the 10th, 50th, and 90th quantiles of habitat productivity as measured by PC1. Panel (b): the approximate ratio of true $S_{MSY}$ relative to estimates assuming semelparity as a function of smolt to adult return rate and annual adult survival ($e^{-M}$; eq. 41) between spawning events.

## 1.J. Supplementary Materials

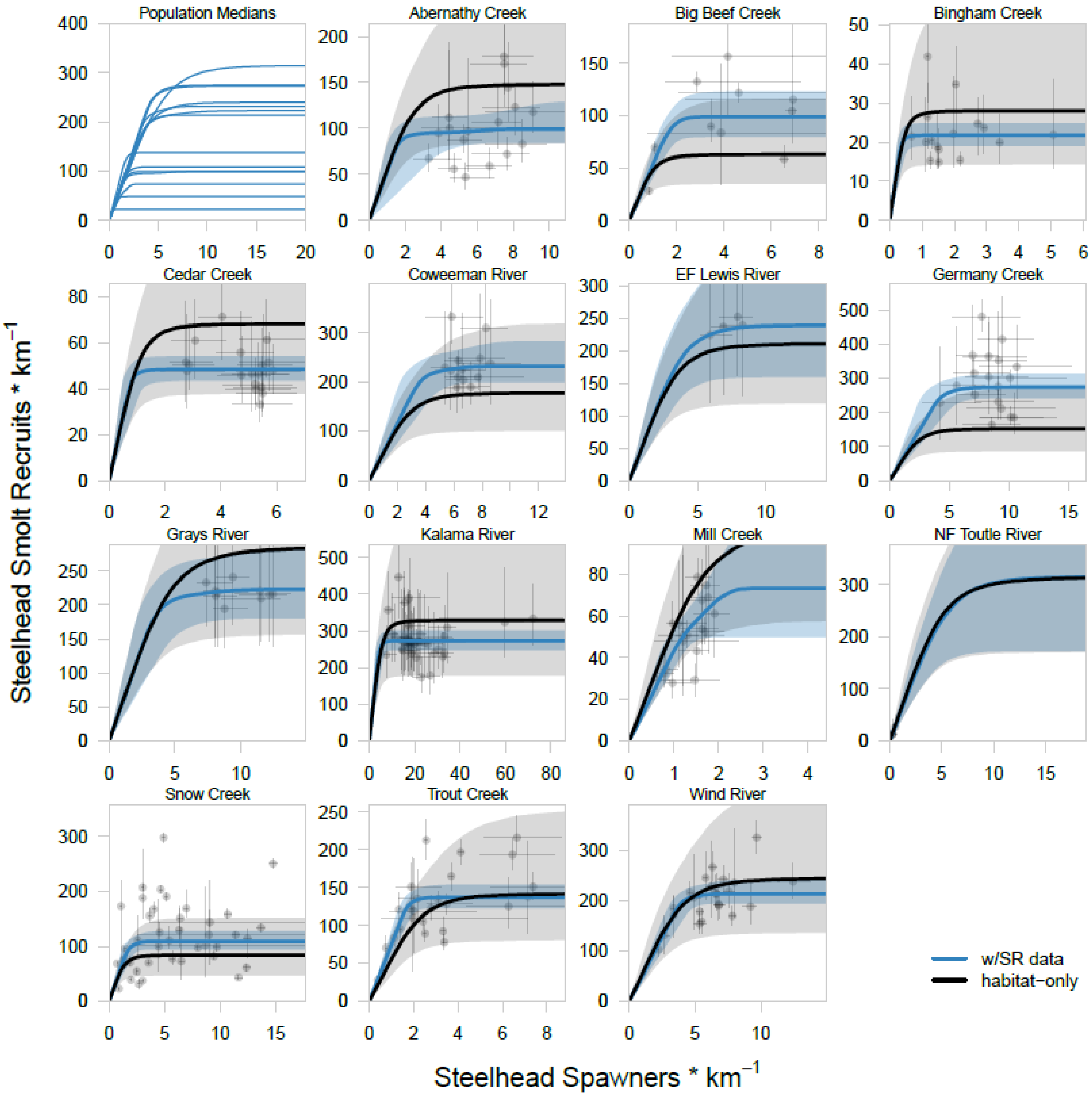


Figure S1. Estimated wild steelhead smolt recruits plotted against estimated brood year spawner abundance for 14 coastal populations in Washington State. Points are medians and grey lines show 90% CI on both axes for the estimates of each spawner-recruit pair. Blue lines are the posterior median estimate of recruitment for the hockey stick model of the stock-recruitment relationship for each population. Black lines are the estimated recruitment based on the posterior predictive distribution (excluding fish data for each population) estimate of productivity and capacity using accessible stream length and PC1. Shading indicates 80% credible intervals. Note that although the hockey stick results in a segmented regression, the predicted recruitment depicted here incorporates uncertainty and covariance in productivity and capacity, which results in the smoothing of the intersection between the two line segments of the Hockey Stick.

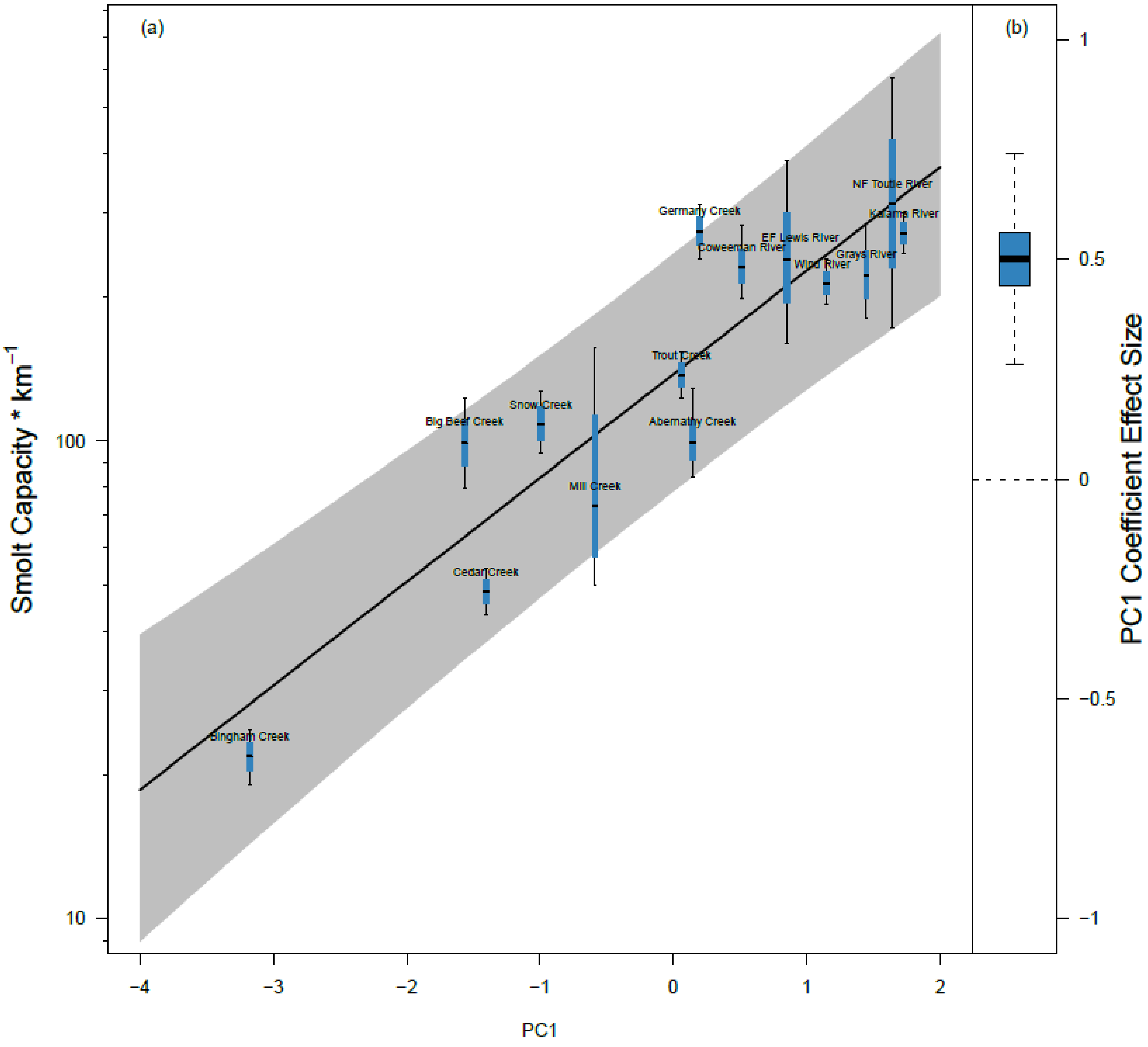


Figure S2. The relationship between estimated smolt capacity ($\alpha S^*$ or $K$) per kilometer of habitat and PC1for the hockey stick model for 14 coastal steelhead populations (a), and the posterior distribution for the coefficient estimating the effect of PC1 on capacity (b). Blue lines and whiskers are medians and 80% credible intervals of the posterior distribution for each population. The black line and grey polygon show the median and 80% credible interval for the posterior predictive distribution for an unknown population. Note log scales on y-axis of (panel a).

## Chapter 2: An integrated model to estimate lower Columbia River coho salmon abundance relative to recovery goals

### 2.A. Abstract

Coho salmon are an important part of temperate and sub-arctic socio-ecological systems of coastal areas around the North Pacific Ocean. Despite their importance to fisheries and their provision of ecosystem services, most populations in the contiguous United States are imperiled. The Lower Columbia coho salmon Evolutionarily Significant Unit (ESU) was listed under the U.S. Endangered Species Act in 2005 as Threatened. At the time of listing, most Lower Columbia populations in Washington State were assumed to number less than 50 fish and total abundance for the Washington portion of the ESU was thought to be less than 1,500 natural origin (hereafter, "wild") coho salmon. However, considerable uncertainty existed in abundance estimates because, at the time, nearly all Washington State populations were not monitored. In 2010 an ESU-wide monitoring program was initiated that included spawning ground surveys and mark-recapture experiments at weirs, which complemented existing trap-and-haul counts of migrating coho at dams and fishery catch estimation programs. Using these data, we developed a multivariate state-space integrated population model to estimate coho salmon population parameters and enable evaluation of population status and trends. The model estimated that coho salmon abundance ranged between 12,348 and 85,910 wild spawners (median = 28,310) from 2010-2022 for Lower Columbia populations in Washington State, and that several populations periodically exceeded their ESA recovery goals but remained below densities seen in healthy populations elsewhere. Interannual variation in abundance was relatively synchronous among

populations (mean pairwise correlation 0.55-0.80), but populations above dams that required trap-and-haul operations exhibited higher interannual variation and less synchrony with the remainder of the ESU. The proportion of hatchery spawners varied among years and populations but was frequently above levels recognized to facilitate fit, locally adapted, viable populations

## 2.B. Introduction

Coho salmon *Oncorhynchus kisutch* are an important species of socio-ecological systems throughout the temperate and sub-arctic coastal areas bordering the North Pacific Ocean in Asia and North America. In common with all Pacific salmon *Oncorhynchus* spp., coho salmon sustain economically and culturally important sport, commercial, and tribal fisheries (Groot and Margolis 1991). Relative to other Pacific salmon species, coho salmon play a particularly important role in stream ecosystems due to their widespread distribution, their long freshwater residence prior to smolting, and their tendency to use smaller streams for spawning than other species (Quinn 2018). Despite their cultural, ecological, and economic importance, many populations in the continental United States are imperiled, and all but those occurring in Puget Sound and the Washington coast are listed as either Threatened or Endangered under the U.S. Endangered Species Act, ESA (Ford 2022).

Effectively managing, conserving, and recovering coho salmon requires that their status and trends be monitored. McElhany (2000) identified four characteristics termed Viable Salmon Population (VSP) parameters important for assessing the health of populations: abundance, productivity, spatial structure, and diversity. These parameters enable identification of trends, assessment of extinction risk, inform conservation efforts, and facilitate assessment of the efficacy of recovery actions. VSP monitoring is therefore essential to the conservation and management of coho populations. Despite the paramount importance of VSP monitoring, many populations either wholly or partially lack rigorous monitoring programs that enable unbiased estimation of these important population parameters (e.g., Crawford and Rumsey 2011).

The Lower Columbia Evolutionarily Significant Unit (ESU), which comprises 24 coho salmon populations from the mouth of the Columbia River upstream to the White Salmon and Hood Rivers in Washington and Oregon states, respectively, was listed as Threatened under the ESA in 2005. This ESU faced grave threats at the time of listing, including widespread destruction of freshwater habitat, historic exploitation rates that exceeded 90%, dams that blocked volitional upstream fish passage on the six of the largest rivers in the ESU (Cowlitz, Lewis, Clackamas, Sandy, White Salmon, and North Fork Toutle), and hatchery programs that had produced many times the abundance of wild populations for decades, likely affecting them through genetic interactions, competition, and other common hatchery impacts (Naish et al. 2007, Myers et al. 2006, Good et al. 2005). In addition, all 17 of the coho populations in Washington State lacked robust VSP monitoring data except for five populations with counts of adults trapped and trucked around fish passage-blocking dams. As a result, it was thought that the total abundance of wild spawners in the ESU at the time of listing numbered less than 1,500 in Washington State (NMFS 2013).

In 2010, the Washington Department of Fish and Wildlife (WDFW) initiated an ESU-wide monitoring program to monitor the VSP parameters of the coho populations in the Washington portion of the Lower Columbia ESU. Monitoring coho salmon populations presents unique challenges due to the widespread distribution of coho salmon, and the timing of their spawning, which peaks coinciding with the period of maximum annual stream flows in lowland areas of Western Washington and Oregon (November-December), complicating observation of live and dead spawners and their nests (redds). The goal for the ESU-wide monitoring program was to develop unbiased estimates of VSP parameters for the populations, with both quantifiable and low (i.e. 15% CV for abundance) uncertainty, which was a novel approach at the time. Previously developed monitoring programs for coho salmon elsewhere in Washington that relied chiefly upon spawning ground survey counts of live spawners, carcasses, and redds, lacked methods to quantify uncertainty and mostly involved opportunistic spatial sampling designs, negating their ability to obtain unbiased abundance estimates. Oregon had made considerable progress developing statistical sampling design-based spawning ground survey methods for the Oregon Coast ESU (Firman and Jacobs 2004). However, this design primarily aimed to obtain unbiased redd estimates for a population but required additional assumptions (e.g., fish-per-redd) to convert redds to spawner abundance and did not directly integrate direct estimates of spawner abundance (i.e., from census counts or mark-recapture) from which estimates of fish-per-redd could be generated with uncertainty.

Integrated population models (IPMs) provide a means to integrate multiple disparate data sources to estimate population parameters via a joint likelihood (Riecke et al., 2019). This approach preserves the uncertainty of each of the data sets and allows all data sets to inform estimates of all parameters jointly and simultaneously. IPMs are easily extended to a multivariate state-space formulation, enabling joint estimation of parameters across multiple state variables (e.g., populations) and across years, which facilitates partitioning of variance into components resulting from biological processes, and random components, a considerable portion of which can result from measurement error. This approach has the benefit of sharing information from more data-rich populations and years with more data-poor populations and years, thereby reducing overall uncertainty and enabling estimation of parameters for variables where little or no data are available.

Our objective was to estimate VSP parameters for the coho salmon populations comprising Washington's portion of the Lower Columbia ESU. Specifically, we developed a multivariate state-space IPM to: 1) quantify spawner abundance and density and compare them with ESA recovery goals and biological reference points developed elsewhere, 2) quantify the contribution of hatchery-origin coho to naturally spawning populations, and 3) understand the role of local vs regional factors in driving variation in abundance.

## 2.C. Methods

### *Study Location*

The Lower Columbia River (LCR) Coho Salmon Evolutionary Significant Unit (ESU) is composed of 24 populations split between Washington and Oregon state (Myers et al. 2006). Our study focused on the 17 coho salmon populations that comprise the Washington State portion of the Lower Columbia River coho salmon ESU (Figure 1, Table 1). Of these, two, the Lower Gorge and Upper Gorge/White Salmon populations contain habitat in both Oregon and Washington. In 2010, the WDFW initiated a program to monitor the VSP parameters of coho salmon populations, and to recover Coded Wire Tags (CWT). This geography of this program spanned the Grays-Chinook Rivers population at the mouth of the Columbia River to Bonneville Dam at river mile 146.1. Prior to 2011 much of the habitat in the Upper Gorge/White Salmon population was blocked to anadromy by Condit Dam. The dam was removed in 2011 (Allen et al. 2016) reopening formerly blocked habitat and enabling coho salmon to establish a population. However, the monitoring program lacked funding for expansion, so we report on all populations except the Upper Gorge population. Within Washington State portion of the lower Columbia River coho salmon ESU, the study area was further divided into 28 sub-populations based on either fisheries management needs or research interest in specific sub-units, or in other cases based on the location of monitoring infrastructure which led to within-population differences across subpopulations in the availability of observational data. Monitoring coho salmon involved collecting numerous data types (Figure 1), including spawning ground surveys, operating weirs, trap-and-haul facilities at dam fish ladders, and using WDFW's catch record card system to estimate catch of pre-spawn adults in a subset of trap-and-haul locations where estimation of adult abundance occurred pre- rather than post-fishery. Details of each data collection program follow.

### *Spawning Distribution Sampling Frame*

A sampling frame to delineate the entire portion of the stream network where spawning could occur (Stevens et al. 2007) was developed using an occupancy model designed to estimate the uppermost extent of coho salmon spawning distribution in each tributary. This approach assumed that distribution was continuous below upper extents downstream to tidewater or the lowermost extent of spawning gravel. The parameters of the occupancy model were estimated based on the methods of Fransen et al. (2006) and Walther et al. (2021) using a probability cutoff for the upper extent of distribution beyond which coho salmon were unlikely to occur. More thorough descriptions of methods are available in Rawding et al. (2010) and Rawding et al. (2014).

### *Generalized Random Tessellation Stratified (GRTS) Survey Sampling Design*

The sampling frame described above was used as the basis for a spatial sampling design for spawning ground surveys developed for 12 of the 17 coho salmon populations in Washington. The Upper Gorge-White Salmon population was excluded due to limited resources, the Mill-Abernathy-Germany (MAG) Creeks, Upper Cowlitz, Cispus, and Tilton River populations were excluded because there were census spawning ground surveys of all habitat in the MAG and census dam counts of pre-spawn adults in the Upper Cowlitz, Cispus, and Tilton River populations. For each population a GRTS sampling design was used to establish a set of random, spatially balanced sample points for coho salmon surveys (Stevens 2002). Reach selection was based on the Lower Columbia River GRTS web-based sampling tool developed by Oregon State University (OSU) through the Pacific Northwest Aquatic Monitoring Partnership (PNAMP) with assistance from Don Stevens (OSU). Reaches, one mile in length, were established based on these points. In a few cases the reach length was less than one mile. This occurred when the GRTS point was in a small tributary less than one mile in overall length or there was an anadromous barrier falls within a mile from the mouth. A three-year rotating panel design was established for each coho salmon population (Firman and Jacobs 2004). In this design about 1/3 of the surveys for the 9-year period are repeated annually, 1/3 are repeated every third year, and new points are chosen each year for the remaining 1/3 of all surveys.

### ***Spawning Ground Surveys***

Redd surveys followed the protocols of Gallagher et al., (2007). Surveyors located the uppermost point in the reach and walked downstream to the coordinates at the bottom end of the reach each week beginning on October 1 and ending on the later of December 31 or when two weeks had passed since the last new redd or live spawner was observed. This period was selected to cover the entire temporal extent of coho spawning. During surveys all identifiable new redds were flagged, and their location (GPS coordinates) recorded. In subsequent surveys, previously flagged redds were inspected to determine if they should be classified as “still visible” or “not visible”. A redd was “still visible” if it would have been observed and identified without the flagging present and was “not visible” if this criterion was not met. These data were collected to allow estimation of the time period redds were visible to surveyors. In addition, all live adult and jack (< 47 cm; T. Buehrens, WDFW, *unpublished data*) salmonids were counted by species and the presence or absence of Floy™ T-bar tag(s) and their color were noted. Salmon were identified as either holding or spawning. A fish was identified as holding if observed in an area not considered spawning habitat, such as pools or large cobble and boulder riffles (Parken et al. 2003), and as spawning if observed on a redd or not classified as a holder. All carcasses that were not totally decomposed were sampled for external tags (Floy T-bar or opercle tags) and biologically sampled for fork length, sex, adipose fin presence (to determine wild or hatchery origin), and condition (extent of decomposition). All fish were sampled for CWT following standard protocols (NWMFT 2001). Spawning success was approximated based on visual estimation of the proportion of eggs retained in dead females, ranging from 100% to 0% success in 25% increments.

### ***Weirs***

A series of partial capture weirs and fish traps was operated throughout the sampling frame for this study (Table 1, Figure 1). These weirs had a diversity of original purposes and designs, but all facilitated the use of mark-recapture techniques wherein the portion of the subpopulation captured and placed upstream of the weir trap could be externally tagged, enabling Petersen mark-recapture estimates of abundance (e.g., Schwarz and Taylor 1988) via visual counts of tagged and untagged live and dead fish upstream of the weirs. Weir operations yielded usable mark-recapture data in Abernathy (two locations), in Delameter, Olequa, Ostrander, Lacamas, and Duncan Creeks. The standard Petersen assumptions applied to these experiments: 1) there was no mark loss, 2) there were no marking effects, 3) all marked and unmarked fish were correctly identified and enumerated, 4) the population was closed, and 5) all fish in the population had the same probability of being tagged or all fish had the same probability of being captured in the second sample; or marked fish mixed uniformly with unmarked fish (Schwarz and Taylor 1998).

Live fish intercepted at each weir in good condition were anesthetized, bio-sampled, double Floy™ tagged (FD 68BC T-bar Anchor tags, Floy Tag & Mfg., Inc. Seattle, WA), given an operculum punch, and then released upstream. Only wild adults (identified by intact adipose fins and lack of CWT) were tagged and passed upstream at some locations and in some years. The number of observable redds constructed per female spawner was estimated by spatially pairing adult abundance above the weirs from mark-recapture operations, redd counts from spawning ground surveys above weir locations, and sex ratio of spawners from biological data. In areas where mark-recapture was not possible (e.g., below weirs), redd counts were translated into spawner abundance by applying sex ratio estimates from carcasses and estimates of the ratio of female spawners to redds generated from above-weir areas to the estimate of total redds (see *Statistical Model* description).

### ***Trap-and-haul***

Several populations have all or a portion of their habitat located above fish-passage blocking dams (Table 1, Figure 1). One-way (NF Toutle) and two-way (Lewis, Cowlitz basins) trap-and-haul programs have been implemented at these locations to enable spawners to use upstream habitat. In one-way trap-and-haul locations, smolts migrated unassisted. Two-way locations require capture and transport of both adults and smolts for wild populations to persist.

Dam counts were used at the three locations with trap-and-haul data, and provided a census count of the number of pre-spawn adults that were passed upstream (Table 1). Locations included the Barrier Dam on the Cowlitz River, the Toutle Fish Collection Facility (TFCF) on the NF Toutle River, and Merwin Dam on the Lewis River. Fish captured at the Cowlitz Barrier Dam were released into each of three upstream populations based on the distinct marks they received as juveniles (hatchery or wild), indicating their population of origin. We made the following key assumptions for the trap-and-haul programs: 1) transported fish spawned in the watershed where they were released (there was no fall back), 2) when fisheries in the Upper Cowlitz, Cispus, and Tilton Rivers occurred, only marked (adipose clipped fish) were harvested in accordance with regulations and there was no illegal harvest, 3) unmarked fish were handled and released in fisheries at the same rate as hatchery marked fish, and sustained 10% catch-and-release mortality, which is higher than recent published estimates of post-release coho salmon mortality in the region (Courter et al. 2023) and may therefore account for some sublethal effects of fishery handling on spawner productivity. Fish captured and transported at these dams were sampled for biological data including, sex, adipose clip presence, CWT presence, and age. Adipose clips and CWT were used in conjunction to identify fish as wild (neither clip nor CWT) or hatchery origin. In general, all hatchery origin coho released in the lower Columbia region in Washington are either adipose clipped or coded wire tagged. WDFW's catch record card system was used to generate estimates of hatchery-origin catch occurring above these dams, which was used to adjust trap-and-haul counts to enable estimation of spawner abundance.

### *Statistical Model Overview*

Coho abundance was estimated using a multivariate autoregressive state-space model. The model was comprised of a biological process model and likelihoods to condition the parameters in the process model. The process model describes the relationships between state variables (parameters) and their co-evolution over time and is intended to represent the biology of the system. The parameters of the process model are informed by observation models that relate the unobserved process model states to observed data using likelihoods. The goal of the model was to use a relatively simple process model to generate abundance estimates which were then conditioned by likelihoods using a more complex observation model that incorporated disparate observation data, therein propagating the various error sources. This approach contrasts with many previous integrated models for salmon (e.g., Fleischman et al. 2013, Buhle et al. 2018) that have involved more complex process models, such as recruitment functions, and multiple life stages, but relatively simple observation models that rely on considerable data pre-processing, such as pre-existing population-scale estimates of abundance and harvest estimates.

For purposes of estimating coho abundance, total adult abundance (wild plus hatchery origin) in each sub-population was defined as the state which would evolve over time, and which through various transformations and relationships with other parameters defined in the process model, could be related to observed data to estimate parameters of interest (e.g., abundance, spawner density, and the proportion of the population that was marked—adipose-clipped or coded wire tag positive). The model was run for spawn years 2010-2022 for 28 sub-population monitoring units comprising 16 of the 17 populations, excluding mainstem areas the mainstem Cowlitz and Toutle Rivers, where spawning is not believed to occur, the mainstem NF Lewis

River, where minimal spawning occurs, and streams comprising the Upper Gorge population, which is not currently monitored (Table 1).

### *Process Model*

Wild and hatchery-origin coho salmon density, redd abundance, and spawner abundance were estimated via process models. Wild spawner density followed a log-normal multivariate random walk with both regionwide-scale process errors and uncorrelated subpopulation-scale process errors with unequal variance. The proportion of hatchery spawners also followed a multivariate random walk with uncorrelated subpopulation-scale process errors with equal variance. Wild spawner abundances were calculated by multiplying wild spawner densities by the length of stream network in the spawning frame for each population. Hathery and total abundance were calculated by applying estimates of the proportion of hatchery spawners to the wild spawner abundance. Observable redd totals for each subpopulation and year were calculated by multiplying total spawner abundance by the proportion of females estimated for each subpopulation-year combination as a logit-normal random effect (yielding an estimate of the number of females spawners) and then multiplying by the observable number of redds per female spawner which was a global parameter estimated across all sub-populations and years. In general, process models used vague or regularizing priors and non-centered parameterizations that decomposed errors into unit-normal random variates which were rescaled by their standard deviations (Table 2).

The logarithm of the density $\lambda$ of wild adult spawners for each sub-population $p$ was assumed to follow a multivariate random walk, where the value in year $y$ was equal to the value in the previous year plus two white noise process error terms, one reflecting regional-scale variation $z_{\eta}$ and the other reflecting sub-population-scale variation $\boldsymbol{z}_{\lambda}$(e.g., Scheuerell et al., 2015). We chose a random walk rather than a lag-3 process (spawners to offspring) due to the short duration of the timeseries and because there was evidence that interannual variation from marine survival has stronger influence on interannual variation in abundance than parental spawner abundance.:

$$\log(\boldsymbol{\lambda}_{y}) = \log(\boldsymbol{\lambda}_{y-1}) + z_{\eta_{y}}\sigma_{\eta} + \mathbf{L}_{\lambda}\boldsymbol{z}_{\lambda_{y}} \quad (1)$$

The regional- and sub-population-scale process errors were estimated based on non-centered and scaled parameterizations (Papaspiliopoulos et al. 2007) to facilitate convergence, where $z_{\eta}$was an annual unit-normal random variate representing regional-scale process variation, and was multiplied by the scale of that variation $\sigma_{\eta}$. Similarly $\boldsymbol{z}_{\lambda}$was a *p*-length annual vector of unit-normal independent random variates describing sub-population process variation that were multiplied by $\mathbf{L}_{\lambda}$, the Cholesky factor of the sub-population process error covariance matrix, whose off-diagonal elements were zero, and whose diagonal elements were the vector of sub-population-specific process error standard deviations $\boldsymbol{\sigma}_{\lambda}$. The vector of population-specific process error standard deviations was estimated as a random effect:

$$\boldsymbol{\sigma}_{\lambda} = e^{\log(\mu_{\lambda}) + \boldsymbol{z}_{\sigma}\sigma_{\lambda_{\sigma}}} \quad (2)$$

where $\mu_{\lambda}$ is the median standard deviation, $\boldsymbol{z}_{\sigma}$ is a *p*-length vector of unit-normal random effects and $\sigma_{\lambda_{\sigma}}$ is the scale of the random effects.

The logit of the proportion of hatchery spawners was assumed to follow a random walk where $\boldsymbol{p}_{H_y}$ was a vector of length $p$ containing the annual subpopulation estimates. The model for $\boldsymbol{p}_{H_y}$, which was assumed to be independent for each subpopulation, does not include a shared process error term because independent anthropogenic factors (e.g. hatchery production in a specific watershed) drive variation in the proportion of hatchery fish:

$$\text{logit}(\boldsymbol{p}_{H_y}) = \text{logit}(\boldsymbol{p}_{H_{y-1}}) + \mathbf{z}_{H_y}\sigma_{\text{H}} \tag{3}$$

where $\boldsymbol{z}_H$ was a $p$-length annual vector of unit-normal independent random variates describing sub-population process variation in the proportion of hatchery spawners that was multiplied by the scale of that process variation $\sigma_{\text{H}}$, which was assumed to be independent of subpopulation.

Wild spawner abundance was the spawner density multiplied by the sample frame length of available habitat in stream kilometers $\boldsymbol{K}$ (eq. 4). Hatchery-origin spawner abundance was the wild spawner abundance multiplied by the odds ratio of hatchery spawners (eq. 5), and total spawner abundance $\boldsymbol{S}_y$ was simply their sum:

$$\boldsymbol{S}_{W_y} = \boldsymbol{\lambda}_y \boldsymbol{K} \tag{4}$$

$$\boldsymbol{S}_{H_y} = \boldsymbol{S}_{W_y}\boldsymbol{p}_{H_y}/(1 - \boldsymbol{p}_{H_y}) \tag{5}$$

$$\boldsymbol{S}_y = \boldsymbol{S}_{H_y} + \boldsymbol{S}_{W_y} \tag{6}$$

The state vector of the number of observable spawning redds (nests) constructed each year $\boldsymbol{R}_y$ for each subpopulation was a function of the total spawner abundance, the proportion of females $\boldsymbol{p}_{F_y}$ that year in each population, and the number of observable redds per female γ, which was a global parameter.:

$$\boldsymbol{R}_y = \boldsymbol{S}_y \boldsymbol{p}_{F_y} \gamma \tag{7}$$

where the logit of $\boldsymbol{p}_{F_y}$ was a non-centered random effect across all years and populations:

$$\text{logit}(\boldsymbol{p}_{\boldsymbol{F}_{\boldsymbol{y}}}) \ = \text{logit}(\mu_F) + \ \boldsymbol{z}_{\boldsymbol{F}_{\boldsymbol{y}}} \sigma_{\mathrm{F}} \tag{8}$$

where $\mu_F$ was the median proportion of females, $z_{F_{p,y}}$was unit-normal random variate and $\sigma_{\mathrm{F}}$ was the scale of among subpopulation and year variation in the proportion of females.

Spawner abundance estimates for subpopulations occurring above dams where no spawning ground surveys were conducted (i.e. trap-and-haul A, B, C, D; Table 1, Figure 1) relied on counts of pre-spawn wild and hatchery-origin adults trapped at the dam and trucked to release sites upstream, where fisheries reduced abundances. Annual fishing exploitation rates for hatchery $\boldsymbol{M}_H$ and wild $\boldsymbol{M}_W$ were used to relate pre-fishery counts at dams with process model state estimates of wild and hatchery-origin spawner abundance from the model. These exploitation rates, like the proportion of females, were estimated as random effects across all subpopulations and years because anthropogenic (e.g., fishery structure) rather than natural processes were believed to contribute to unpredictable among-subpopulation and -year variation:

$$\text{logit}(\boldsymbol{M}_{H_y}) = \text{logit}(\mu_{\boldsymbol{M}}) + \mathbf{z}_{M_y} \sigma_M \tag{9}$$

where $\mu_{\boldsymbol{M}}$ was $p$-length vector of medians above dam exploitation rates, $\mathbf{z}_{M_y}$was a $p$-length annual vector of unit-normal independent random variates describing sub-population process variation in the above-dam fishing exploitation rate, that was multiplied by the scale of that process variation $\sigma_{\mathrm{M}}$ which was assumed to be independent of population. The fishing mortality rate on wild fish was the hatchery exploitation rate multiplied by 0.1, which assumed fisheries captured the same proportion of the population of wild adults as hatchery adults but released them with 10% release mortality.

### *Observation Models*

The observed number of hatchery-origin carcasses for each year and subpopulation was assumed to follow a binomial distribution with $s_{HW_{p,y}}$ equal to the number of carcasses where origin was identifiable, and the year- and subpopulation-specific proportion of hatchery-origin $p_{H_{p,y}}$ defined in the process model:

$$s_{H_{p,y}} \sim binomial(s_{HW_{p,y}}, p_{H_{p,y}}) \tag{10}$$

Similarly, the observed number of female carcasses for each year and subpopulation was assumed to follow a binomial distribution with $s_{FM_{p,y}}$ equal to the number of carcasses where sex was identifiable, and the year- and subpopulation-specific proportion of female spawners $p_{F_{p,y}}$ defined in the process model:

$$s_{F_{p,y}} \sim binomial(s_{FM_{p,y}}, p_{F_{p,y}}) \tag{11}$$

For a subset of subpopulations and years, a portion of upstream migrating spawners was captured at weirs and released upstream after tagging $n1_{p,y}$ with externally visible double Floy™ tags and a permanent operculum holepunch (“marks”), and a subset of these $m_{p,y}$ was either resighted alive or recovered as carcasses during upstream spawning ground surveys (“recaptures”). These recaptures were also a subset of all recovered carcasses and sighted live adults $C_{p,y}$ (“captures”):

$$m_{p,y} \sim binomial(C_{p,y}, p_{M_{p,y}}) \quad (12)$$

$$n1_{p,y} \sim Poisson(S_{p,y} p_{M_{p,y}}) \quad (13)$$

The most common data type was counts of redds from spawning ground surveys. To compare redd estimates with observations made using multiple study designs, it was necessary to estimate the proportion of redds that occurred in areas with non-representative “index” surveys $p_I$, versus those occurring in representative surveys selected with Generalized Random Tessellation Stratified (GRTS) $(1 - p_I)$. For subpopulations where the entire spawning distribution was surveyed using “index” surveys, $p_I$ was set to 1, and conversely to 0 when solely GRTS surveys were used. For subpopulations with mixed study designs, the parameter was estimated each year it was unknown. The parameter was then used to relate the sum of index counts for that subpopulation and year $r_{I_{p,y}}$ to the total redd count $R_{p,y}$:

$$r_{I_{p,y}} \sim Poisson(R_{p,y} p_{I_{p,y}}) \quad (14)$$

Using this parameterization, index redd counts were used to inform estimates of total redds in the index reaches and were not assumed to be representative of densities elsewhere. In contrast, GRTS redd survey reaches were selected to be representative of the un-surveyed areas, meaning that it was reasonable to apply the mean and variance of redd density in the GRTS reaches both to the GRTS survey reaches themselves, as well as the unsampled habitat. The expected redd density (redds·km$^{-1}$) in particular GRTS reaches and un-surveyed areas $\mu_{R_{p,y}}$ was calculated as:

$$\mu_{R_{p,y}} = \frac{R_{p,y}\left(1-p_{I_{p,y}}\right)}{K_G + K_U} \tag{15}$$

which was essentially the subpopulation redds estimated to occur in GRTS reaches and un-surveyed areas divided by the stream kilometers comprising GRTS reaches $K_G$ and un-surveyed areas $K_U$. This density could then be compared directly with GRTS redd counts in a particular reach, year, and subpopulation $r_{p,y,i}$:

$$r_{p,y,i} \sim negative\ binomial(\mu_{R_{p,y}} K_i,\ \sigma_{D_p}^{-2}) \tag{16}$$

where $K_i$ was the length of the GRTS reach and $\sigma_{D_p}^{-2}$ was the sub-population-specific negative binomial dispersion parameter, which was parameterized as a standard deviation and raised to the power -2 (transformed to a precision) for use in the negative binomial. The vector of population-specific dispersion parameters was estimated as a random effect:

$$\boldsymbol{\sigma}_D^{-2} = e^{\log(\mu_D) + \boldsymbol{z_D} \sigma_{D_\sigma}} \tag{17}$$

where $\mu_D$ is the mean negative binomial standard deviation (dispersion$^{-2}$) $\boldsymbol{z_D}$ is a $p$-length vector of unit-normal random effects and $\sigma_{D_\sigma}$ is the scale of the random effects.

The trap-and-haul census counts of pre-spawn hatchery $T_{H_{p,y}}$ and wild $T_{W_{p,y}}$ adults for subpopulations occurring above dams were related to model estimates of total wild $S_{W_{p,y}}$ and hatchery $S_{H_{p,y}}$ spawners, adjusting for harvest occurring after release above the dam:

$$T_{H_{p,y}} \sim Poisson\left(\frac{S_{H_{p,y}}}{1-M_{H_{p,y}}}\right) \tag{18}$$

$$T_{W_{p,y}} \sim Poisson(\frac{S_{W_{p,y}}}{1-0.1M_{H_{p,y}}}) \qquad (19)$$

In instances where no fisheries occurred above the dam the harvest mortality term was removed from eqs. 18-19. Finally, catch estimates for hatchery catch $c_{p,y}$ and their associated sampling variance estimates $\sigma^2_{C\,p,y}$ generated from the WDFW's catch record card system were related to the model estimates of catch:

$$c_{p,y} \sim lognormal(\log\,(S_{H_{p,y}}\frac{M_{H_{p,y}}}{1-M_{H_{p,y}}}) - \sigma^2_{C\,p,y}/2, \sigma_{c_{p,y}}) \qquad (20)$$

where $c_{p,y}$ is assumed to the MLE of a lognormal estimator, and thus the mean estimate of catch from this model $S_{H_{p,y}}\frac{M_{H_{p,y}}}{1-M_{H_{p,y}}}$ must be adjusted by subtracting $\sigma^2_{C\,p,y}/2$.

### *Priors*

We used minimally informative priors in almost all cases to let the data dominate the posterior (Table 2). Standard deviations were given half normal and half Cauchy distributions to induce regularization. The only notably specific prior was the use of an empirical Bayes hyper-mean to initiate the log spawner density initial states. However, that prior was still very vague to let the data dominate the posterior.

### *Summarizing Results at the Population Scale*

State variables in the model were estimated at the subpopulation scale, requiring post-model calculation of derived quantities at the population scale. To estimate parameters at the NOAA-designated population scale (Table 1), subpopulation abundances were summed on a draw-by-draw basis to preserve uncertainty. The proportion of hatchery spawners, wild and hatchery-origin spawner densities, and interannual correlations in spawner abundances were calculated by first summing subpopulation-scale abundances and then calculating the corresponding population-scale metrics. Densities at the watershed scale were calculated by dividing wild and hatchery-origin abundance by the spawning frame kilometers. Densities were then compared with results of a meta-analysis by Bradford (2000) which estimated the number of spawners necessary to saturate smolt production ($N$*) using a Hockey Stick model.

### *Model Fitting and Validation*

The model was implemented in a Bayesian framework and was fit using Stan software (Carpenter et al., 2017). Stan samples the posterior distribution using the "No-U-Turn" algorithm, which is a form of Hamiltonion Monte-Carlo (Hoffman and Gelman 2014). This algorithm optimizes sampling to reduce autocorrelation among samples of the posterior, enabling more efficient estimation. The sampling had a "warm up" of 1,000 draws during which Stan tunes the sampler, and which consequently, are not used for inference. Following warm up, 3,000 additional samples were drawn from each of 4 chains, leading to 12,000 samples from the posterior. The chains were not thinned because autocorrelation is minimized by the No-U-Turn algorithm, and thinning wastes information. Model convergence was determined from visual assessment of trace plots for chain mixing, evaluation of the Brook-Gelman-Rubin statistic ($\hat{R}$; Brooks et al., 2011) and the effective sample size. Goodness of fit was assessed by generating posterior predictive distributions for all likelihoods and calculating the proportion of observations falling within nominal 95% posterior predictive intervals.

## 2.D. Results

### *Model Diagnostics*

Evaluation of model convergence and fit to the data suggested that the model adequately converged on the posterior distribution of the model parameters and did not exhibit lack of fit to the data. The minimum effective sample size was 1,209 and the maximum $\widehat{R}$ statistic was 1.005. Posterior predictive checks indicated that >= 95% of observations fell within the 95% posterior predictive interval for all likelihoods (Figures S1-9).

### *Spawner Abundance and Density*

Estimates of wild spawner abundance across Washington's populations within the Lower Columbia River coho salmon ESU varied considerably both in magnitude and relative to ESA recovery goals. Wild populations ranged in abundance from the largest population, the Lower Cowlitz (median = 4,365 spawners; interannual range = 2,474-24,544) to the smallest, the Kalama (median = 75, range 31-268) (Table 3, Figure 2). Population abundance consistently exceeded ESA recovery goals for the Coweeman and NF Lewis populations throughout the 13-year study period. The NF and SF Toutle populations, Upper Cowlitz, Cispus, Lower Cowlitz, and EF Lewis populations all exceeded their goals in around half of the years studied. In contrast, all populations in the Coastal Major Population Group (Elochoman-Skamokawa, Grays-Chinook, and Mill-Abernathy-Germany-Coal) achieved the recovery goal in only 2014, and the Kalama, Lower Gorge, and Washougal never achieved their recovery goals. At an ESU level, the aggregate abundance of Washington populations (median = 28,311, range = 12,348-85,910, excluding the Upper Gorge-White Salmon) exceeded the sum of the corresponding population-level abundance recovery goals (23,700) in 10 of 13 years even though the Lower Gorge population estimate only included spawners in Washington (the goal for that population includes spawning habitat in Oregon).

Wild spawner densities ranged from a high in the Coweeman (median = 36.7 spawners/km, range = 12.96-72.00) to a low in the Kalama (median = 1.45 spawners/km, range = 0.61-5.19). Applying a proportion of 0.47 females to the Bradford et al. (2000) estimates of N* (the density of spawners necessary to fully seed freshwater habitat using a hockey stick recruitment model), none of the Lower Columbia populations had a median abundance exceeding the median of the Bradford et al. (2000) predictive distribution (19 female spawners/km or 40 total spawners/km); however, three populations had densities of wild spawners exceeding this level in at least one year (Coweeman, SF Toutle, and Tilton). Additionally, eight out of 15 populations had median densities exceeding the minimum population-specific N* density estimated in Bradford et al. (2000) of 8.18 spawners, and only the Kalama did not exceed this density in at least one year. No populations exceeded the maximum population-specific N* density estimate from Bradford et al., (2000) or 92.3 spawners/km in any year (Table 3, Figure 3).

### *Proportion, Abundance, and Density of Hatchery Spawners*

The abundance of hatchery-origin coho salmon spawners ranged from a high in the Upper Cowlitz and Cispus (median = 8,861, range = 941-22,620), to a low in Salmon Creek (median = 89, range = 24-273). Nearly all populations had median hatchery spawner abundances in the hundreds except Salmon Creek and the three watersheds above dams with trap-and-haul reintroduction programs where median abundances were several thousand or more (Upper Cowlitz and Cispus, NF Lewis, Tilton) (Table 3, Figure 2). Hatchery spawner densities generally ranged from one to ten spawners/km except for Lower Cowlitz and Salmon Creek (generally <1 spawner/km) and Upper Cowlitz and Cispus, NF Lewis, Tilton, and Kalama, where densities were often > 10 spawners/km (Figure 3).

Asynchronous variation of hatchery spawner abundance relative to wild spawner abundance within and among populations led to considerable temporal variability in the proportion of hatchery origin spawners (pHOS) without a clear regional trend. Two populations had interannual medians < 10% (Salmon Creek, EF Lewis), five populations had interannual medians between 10 and 20% (Lower Cowlitz, SF Toutle, Mill-Abernathy-Germany, Coweeman, and Lower Gorge), two had medians in the 20-40% range (NF Toutle, Elochoman-Skamakowa), and six populations had interannual medians >50% (Washougal, Grays-Chinook, NF Lewis, Tilton, Upper Cowlitz-Cispus, and Kalama). At an ESU level, the aggregated interannual median pHOS of Washington populations (excluding the Upper Gorge-White Salmon) was 46% (range 29-57%) (Table 3, Figure 4).

### ***Proportion of Females and Redds per Female***

The proportion of females and redds per female were important parameters for relating expanded redd totals to spawner abundance. The proportion of females was treated as a global random effect across sub-populations and years. The posterior median for the random effect median was 47.7% female, though individual subpopulation posterior median estimates ranged from 45-56% female among years (Figure 5a). The model estimated posterior median ratio of total observable redds per female spawners to be 0.43 (95% CI 0.39-0.48) (Figure 5b). Conditional on the assumption that each spawner constructed one redd, this may also be interpreted as the observer efficiency for redds using the seven-day survey frequency in this study.

### ***Local vs. Regional-Scale Contribution to Abundance Trends***

To examine the contribution of local versus regional influence on coho salmon abundance trends, the process model decomposed interannual variation in natural origin spawner abundance into three components: shared regional variation, local (subpopulation) variation, and random variation. Local and regional process variation were estimate (eq. 1), whereas random variation was filtered in the likelihoods via the state-space model construction. Log-transformed scaled regionwide trends, calculated as the cumulative sum of the scaled region-wide process errors ($z_{\eta_y}$), show a slight downward trend drop from 2010 to 2011, and increase from 2011 to the study period high in 2014, followed by a sharp drop to the study period low in 2015, and a gradual recovery to above average in the final three years of the study from 2020-2022 (Figure 6a). Log-transformed scaled subpopulation level trends were calculated in the same manner as the regional trends but using subpopulation scaled process errors, where cumulative sums of $z_{\lambda_y}$ were calculated independently for each population across years. These showed a great deal of variability, with some populations increasing, others decreasing, and others fluctuating (Figure 6b). The sub-population process error standard deviations varied tremendously, from a low of 0.2

(Salmon Creek) to a high of 2.4 (Upper Cowlitz and Cispus), though all sub-populations without trap-and-haul had posterior median values below 1.0 (Figure 6c). The proportion of process variation explained by regional-scale variation ranged from a high of 88% (Salmon Creek) down to <10% for the Upper Cowlitz and Cispus (Figure 6d). Although pairwise correlations in wild spawner abundance for all populations were positive (Figure 6e), mean pairwise correlations ranged from 55% (Upper Cowlitz and Cispus) to greater than 80% (Salmon Creek).

## 2.E. Discussion

### *Abundance of Wild Coho Relative to ESA Recovery Goals*

Lower Columbia River coho salmon populations in Washington State were thought to be critically depressed in the decade prior to, and immediately following, listing under the Endangered Species Act in 2005. This was reiterated numerous times by NOAA Fisheries, which concluded in their 2000 and 2005 status reviews that most populations in the Lower Columbia River ESU other than those in the Sandy and Clackamas Rivers in Oregon were extirpated (NOAA 2005), though the 2005 status review noted that wild coho salmon abundance was not monitored for nearly all ESU populations in Washington State. Similarly, the NOAA recovery plan for Lower Columbia River coho salmon adopted in 2013 concluded that wild spawner abundances were at levels consistent with “very low” viability for all the populations located wholly or partially in Washington State and estimated that the combined abundance of these populations was less than 1,450 wild spawners, with all but two populations exhibiting “current” abundances less than 50 spawners (Table 6-4, NOAA 2013). Recent status reviews have started to recognize that abundances were higher than previously thought (based on unpublished results of the work described herein) (Ford 2022).

Although there is little information to rigorously estimate wild spawner abundance in Washington's portion of the Lower Columbia River coho salmon ESU prior to 2010, results of the monitoring program since then paint a starkly more optimistic picture of coho salmon status. All of Washington's populations had median abundances greater than 50 spawners, with two populations (Coweeman and NF Lewis) above their abundance recovery goals in every year of monitoring, six achieving their recovery goals in about half of years, another three achieving the goals in one year, with only three monitored populations failing to achieve their abundance recovery goals at least once. At an ESU scale, abundance ranged from over 12,000 to 85,000 wild spawners, which was approximately 8-59 times the "current" abundance identified for the ESU in the recovery plan in 2013 (NOAA 2013). Washington's populations have likely increased in abundance in recent decades, particularly due to reductions in exploitation rates, which averaged 70-90% for decades prior to the early 1990s and have subsequently averaged 10-30% (Ford 2022). Reintroduction efforts in the Cowlitz and Lewis Basins above dams have also increased natural-origin abundance in recent years. However, in the absence of robust monitoring data, several successive NOAA status reviews may have considerably underestimated the abundance of extant natural-origin populations. Regardless, we estimate that current spawner abundances are much higher than understood prior to the initiation of the contemporary monitoring program.

### ***Density Relative to Expectations of Fully Seeded Rivers***

Spawner densities in Washington's populations remain mostly below those previously identified to fully seed habitat despite abundances being far greater than assumed at the time of ESA listing and in some cases exceeding ESA recovery goals. Bradford et al. (2000), in a meta-analysis of 14 Pacific Northwest coho populations in small and intensively monitored watersheds from Oregon to British Columbia, estimated a mean of 19 female spawners per kilometer (sd = 13, range 4 to 44) would achieve full seeding using a hockey stick recruitment model (Barrowman and Myers 2000), corresponding to a mean of 40.4 total spawners per kilometer assuming the lower Columbia average of 47% females. It is possible that many lower Columbia populations remain under-seeded when only considering wild spawners and ignoring the reproductive output of hatchery spawners.

Alternatively, it is possible that seeding densities from Bradford et al., (2000) are higher than needed to seed Washington's lower Columbia populations. This could result from poor condition of current habitat in our study. Our monitoring included many small streams but also larger watersheds (kilometers of habitat range: 51.58 – 640 km summed by population; Table 1) that may naturally have a lower seeding density per kilometer than the generally small watersheds used in Bradford's estimates (kilometers of habitat range: 1.4-92; Table 1, Bradford 2000) due to the specialization of coho salmon in using smaller streams and off-channel areas in larger rivers (e.g., Quinn 2005). Additionally, watersheds in the Lower Columbia are generally more developed and have undergone more land use conversion (e.g., forest and wetland to silviculture, agriculture and urban development) than those considered by Bradford (2000), which could also have reduced seeding capacities in our study relative to theirs. Finally, it is also likely that available habitat quantity was imperfectly measured in some cases in our study, leading to imprecision in spawner densities. For example, our occupancy model identified considerable habitat in the upper Grays River watershed above a waterfall that appears to severely limit coho salmon passage in many years, possibly explaining low densities in that watershed.

### ***pHOS Relative to Recovery Targets***

Hatchery-origin coho salmon comprised a variable but considerable proportion of spawners in most populations. Hatchery spawners can provide ecological benefits by contributing to the reproductive output of naturally spawning populations, but they can also pose risks by passing on maladaptive or non-native genes that have been selected for in the domestic environment or by competition with wild spawners (Naish et al., 2007, Anderson et al., 2020). As a result, some recovery planning processes have identified targets for the maximum proportion of hatchery origin spawners in naturally spawning populations, as a risk containment measure. For example, the Hatchery Scientific Review Group used quantitative genetic models developed by Ford et al. (2002) to identify targets for the maximum proportion of hatchery-origin spawners that differed depending on the hatchery stock in use and the viability level desired for the population they were spawning in. These include 30% pHOS for “integrated” hatchery spawners that are derived from local wild populations and incorporate wild spawners into broodstock each year, when spawning in populations targeted for high viability and 50% for populations targeted for medium viability.  For segregated hatchery programs that either use non-local broodstocks or solely hatchery origin adults in the broodstock pHOS goals identified by the HSRG are 5% or 10% for populations targeted for high or medium viability, respectively (Paquet et al. 2011;

HSRG 2004). The Mitchell Act Biological Opinion adopted similar thresholds of 10% or 30% depending on whether the hatchery program was segregated or integrated (NOAA 2017).

Evaluating programs relative to pHOS standards was challenging because some programs changed broodstock sources (and thus applicable HSRG standards) over the years, and Mitchell Act Biological Opinion (NOAA 2017) did not begin immediately. However, only two populations had interannual median pHOS below 10%, six had medians below 30%, one was between 30% and 50%, and six had had interannual medians above 50%. This indicates that many populations had hatchery spawner proportions above levels needed for high fitness and population viability (Paquet et al., 2011) during this study.

### *Local vs Regional Variation*

Asynchrony in salmon populations is often credited with increasing metapopulation and fisheries stability and buffering risk (e.g., Schindler et al. 2010). Such asynchrony often arises from local variability in habitat and life history diversity, wherein inter-annual or greater temporal-scale variability in productivity differentially affects different habitats and life histories (Hilborn et al. 2003). In this case, high synchrony would likely reduce resilience as populations seem to wax and wane coherently.

Synchrony among Lower Columbia coho salmon populations appeared to be relatively high and was higher in populations not subject to anthropogenic alterations in fish passage (i.e., dams). In our study, the subpopulation-level process variation in wild spawner abundance, the proportion of process variance explained by regional vs. local factors, and the pairwise correlations in abundance all appeared to differ between populations impacted (or not) by dams or weirs. The subpopulations that occurred above dams requiring trap-and-haul (Upper Cowlitz, Cispus, NF Lewis, Tilton) had the three highest subpopulation levels of process variation and were all in the top six among 28 subpopulations (Figure 6c). These subpopulations also had the lowest proportions of process variation in abundance explained by regional trends, and at the population-scale, except for the NF Toutle, which is stabilized by the Green River subpopulation, had among the lowest pairwise correlations in wild abundance (Figure 6e an 6f). In common with subpopulations above dams with trap-and-haul data, subpopulations above weirs also had among the highest subpopulation process error variation (Figure 6c), and lowest proportion of variation explained by regional factors (Figure 6d). This could be related to the effects of weirs on spawner distribution (e.g., Wilson and Buehrens 2024), which varies in strength among years depending on the frequency and magnitude of freshets that facilitate unimpeded upstream passage by submerging weirs. Years with few freshets could result in fewer spawners reaching upstream areas thereby causing elevated subpopulation variability in wild abundance.

The role of life history, habitat conditions, and marine survival must also be considered to interpret the level of synchrony among coho salmon populations. A major mechanism causing asynchrony in salmon populations is variable age structure, including variable age at ocean entry and variable age composition at adult return (Schindler et al. 2010). The lack of age diversity in coho salmon, almost certainly explains some of the considerable synchrony among Lower Columbia River populations. Approximately 95% of wild coho salmon aged by WDFW between 2010 and 2022 in the Lower Columbia EUS were total age three and had spent one winter at sea and three percent were total age 2 and had spent no winters at sea (jacks) (T. Buehrens, WDFW, unpublished data). This means all other life histories (e.g., two sea winter and zero or two freshwater winter coho) comprised less than 2% of returning adults, and essentially all females, which exert greater influence on population dynamics than males, were age three. The predominance of age three fish is likely both natural for this portion of the coho range and could have been amplified through hatchery practices and loss of habitat diversity. It also poses a risk given that variability in ocean survival may increase as a result of climate change, which could negatively impact some brood years and have persistent effects due to relatively independent dynamics of each three-year brood cycle. Ultimately, resilience would be improved if the effects of dams and weirs on subpopulation variability were reduced, and life history diversity could be increased through habitat restoration and a reduction in hatchery influence.

### ***Benefits and Limitations***

The development of a multivariate state-space integrated model successfully enabled estimation of key VSP parameters for lower Columbia coho salmon populations. The integrated model enabled joint estimation of parameters as a function of diverse data sources (Riecke et al., 2019), as well as sharing of information across populations and years. This was particularly important given that the study design did not collect all data types for all populations or in all years and would have otherwise been insufficient to estimate key parameters for populations lacking certain data (e.g., populations with only spawning ground survey data where redds per female was unknown). In these populations the more traditional method of using exogenous estimates of redds per fish would have been required, whereas that parameter was jointly estimated in our model as a function of the data collected for all populations.

Despite all the benefits of our integrated model approach, several limitations remain. For example, we did not report productivity of populations. That was, in part, because hatchery spawners comprised considerable portions of many populations and their reproductive contributions were unknown but possibly less than that of wild spawners (Anderson et al. 2020). Thus, calculating productivity by dividing wild spawner abundance by total spawner abundance at lag three could result in negatively biased estimates of productivity if hatchery spawners were less productive than wild spawners. Future work could revise the process model to include more demographic detail (e.g., Buhle 2018) and even estimate whether hatchery and wild spawners had different productivity, therein enabling unbiased estimates of this VSP parameter.

Another limitation of our work was the precision of the parameter estimates produced. Crawford and Rumsey (2011) recommended a coefficient of variation of 15% for VSP monitoring programs for ESA-listed salmon and steelhead. Our population-scale estimates of wild spawner abundance had mean annual CVs that ranged from 2%-53% (median population interannual mean = 32%), indicating that despite the sharing of information among populations and years facilitated by the multivariate state space model, precision remained low and did not meet NOAA targets. Future enhancements such as reanalyzing redd count data using a spatiotemporal model (e.g., Hocking et al. 2018) and adding live spawner area-under the curve estimates (e.g., Parsons and Skalski 2010, Hilborn et al. 1999) as an additional likelihood contribution are two possible solutions that may reduce uncertainty in VSP parameter estimates.

### *Conclusions*

The development of an ESU-wide coho salmon monitoring program in Washington State has greatly improved understanding of the status of coho salmon populations in the Lower Columbia River ESU. An integrated model enabled us to leverage disparate but common data sources such as spawning ground survey counts, mark-recapture data, dam and trap counts, and harvest estimates to generate annual abundance estimates for all monitored lower Columbia populations in Washington. Contrary to understanding at the time of ESA listing in the early 2000s, Washington's wild lower Columbia coho salmon number in the tens of thousands and many populations are above or occasionally above their recovery goals, meaning their abundance has either increased substantially since listing or was underestimated at the time of listing. While there is potential to further refine and improve the monitoring design used here, application of the methods presented herein would likely benefit monitoring salmon populations elsewhere in the Pacific Northwest where similar data sources exist and extant methods are not able to leverage existing data jointly and propagate uncertainty.

## 2.F. References

Allen, M. B., R. O. Engle, J.S. Zendt, F. C. Shrier, J. T. Wilson and P. J. Connolly. 2016 Salmon and Steelhead in the White Salmon River after the Removal of Condit Dam–Planning Efforts and Recolonization Results, Fisheries, 41:4, 190-203 http://dx.doi.org/10.1080/03632415.2016.1150839

Anderson, J.H., K. I. Warheit, B. E. Craig, T. R. Seamons, and A. H. Haukenes. 2020. A review of hatchery reform science in Washington State. Final report to the Washington Fish and Wildlife Commission. Olympia, WA: Washington Department of Fish and Wildlife.

Barrowman, N. J., and R. A. Myers. 2000. Still more spawner-recruitment curves: the hockey stick and its generalizations. Can. J. Aquat. Fish. Sci., 57: 665-676.

Bradford, M. J., R. A. Myers, and J. R. Irvine. 2000. Reference points for coho salmon (Oncorhynchus kisutch) harvest rates and escapement goals based on freshwater production. Can. J. Fish. Aquat. Sci. 57: 677–686.

Brooks, S., A. Gelman, G. Jones, & X-L Meng (Eds). 2011. Handbook of Markov Chain Monte Carlo. CRC Press.

Buhle, E. R., M. D. Scheuerell, T. D. Cooney, M. J. Ford, R. W. Zabel, and J. T. Thorson. 2018. Using Integrated Population Models to Evaluate Fishery and Environmental Impacts on Pacific Salmon Viability. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-NWFSC-140. DOI: 10.7289/V5/TMNWFSC-140

Burnham, K. P. and D. R. Anderson. 2002. Model selection and inference: a practical information theoretical approach. New York, Springer-Verlag.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell. 2017. Stan: A Probabilistic Programming Language. J. Stat. Softw. 76(1): 1-32.

Crawford, B.A. and S. Rumsey. 2011. Guidance for monitoring recovery of Pacific Northwest salmon and steelhead listed under the federal Endangered Species Act. NOAA-Fisheries, Portland,OR. 125 pp.

Courter, I., I., T. Buehrens, M. Roes, T. E. Blackman, B. Briscoe, and S. Gibbs. 2023. Influence of angling methods and terminal tackle on survival of salmon and steelhead caught and released in the Cowlitz River, Washington. Fisheries Research, 268: 106848.

Firman, J.C., and S.E. Jacobs. 2004. A Survey Design for Integrated Monitoring of Salmonids. Oregon Dept of Fish and Wildlife. http://oregonstate.edu/Dept/ODFW/spawn/pdf%20files/reports/emappaper.pdf 13p.

Fleischman, S. J., M. J. Catalano, R. A. Clark, and D. R. Bernard. 2013. An age-structured state-space stock–recruit model for Pacific salmon (Oncorhynchus spp.). Can. J. Fish. Aquat. Sci. 70: 401–414.

Ford, M. J. 2002. Selection in captivity during supportive breeding may reduce fitness in the wild. Conservation Biology, 16(3), 815–825. https:// doi. org/ 10. 1046/j. 1523-1739.2002. 00257. X

Ford, M. J., editor. 2022. Biological Viability Assessment Update for Pacific Salmon and Steelhead Listed Under the Endangered Species Act: Pacific Northwest. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-NWFSC-171. //doi.org/10.25923/kq2n-ke70

Fransen, B., S. Duke, L. McWethy, J. Walter, and R. Bilby. 2006. A logistic regression model for predicting the upper extent of fish occurrence based on geographic information systems data. North American Journal of Fish Management 26:960-975.

Gallagher, S.P., P.K.J. Hahn, and D.H. Johnson. 2007. Redd Counts. Pages 197-234 in D. H. Johnson, B. M. Shrier, J. S. O'Neal, J. A. Knutzen, X. Augerot, T. A. O-Neil, and T. N. Pearsons, editors. Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations. American Fisheries Society, Bethesda, Maryland.

Good, T.P., R.S. Waples, and P. Adams (Eds). 2005. Updated status of federally listed ESUs of West Coast salmon and steelhead. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-66, 598 p.

Groot, C., & L. Margolis. 1991. Pacific salmon life histories. University of British Columbia Press, Vancouver, Canada. ISBN 0-7748-0359-2.

Hatchery Scientific Review Group. (2004). Hatchery reform: Principles and recommendations of the HSRG. Hatchery Scientific Review Group.

Hilborn, R., B. G. Bue, and S. Sharr. 1999. Estimating spawning escapements from periodic counts: a comparison of methods. 1999. Can. J. Fish. Aq. Sci. 56(5):888-896.

Hilborn, R., T. P. Quinn, D. E. Schindler, and D. E. Rogers. 2003. Biocomplexity and fisheries sustainability. PNAS, 100 (11): 6564-6568.

Hocking, D. J., J. T. Thorsen, K. O'Neil, and B. H. Letcher. 2018. A geostatistical state-space model of animal densities for stream networks. Ecological Applications, 28 (7): 1782–1796.

Hoffman, M. D., & A. Gelman. 2014. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research, 15, 1351-1381.

McElhany, P., M. H. Ruckelshaus, et al. 2000. Viable salmonid populations and the recovery of evolutionary significant units. NMFS-NWFSC. Tech. Memo 42. 156 pp.

Myers, J., C. Busack, D. Rawding, A. Marshall, D. Teel, D.M. Van Doornik, and M.T. Maher. 2006. Historical population structure of Pacific salmonids in the Willamette River and lower Columbia River basins. U.S. Dept. Commer., NOAA Tech. Memo.NMFS-NWFSC-73, 311 p.

Naish, K.A., J.E. Taylor, 3rd, P.S. Levin, T.P. Quinn, J.R. Winton, D. Huppert, et al. 2007. An evaluation of the effects of conservation and fishery enhancement hatcheries on wild populations of salmon. Advances in Marine Biology, 53, 61–194.

National Marine Fisheries Service. 2013. ESA Recovery Plan for Lower Columbia River Coho Salmon, Lower Columbia River Chinook Salmon, Columbia River Chum Salmon, and Lower Columbia River Steelhead. 503p.

National Marine Fisheries Service. 2022. 2022 5-Year Review: Summary & Evaluation of Oregon Coast Coho Salmon. DOI: https://doi.org/10.25923/95k5-p494.

National Marine Fisheries Service. 2023. 5-year review: summary and evaluation of central California coast Coho Salmon. 94pp. https://doi.org/10.25923/gx95-hz48.

NOAA. 2017. Endangered Species Act (ESA) Section 7(a)(2) Biological Opinion and Magnuson-Stevens Fishery Conservation and Management Act Essential Fish Habitat (EFH) Consultation. NOAA's National Marine Fisheries Service's implementation of the Mitchell Act Final Environmental Impact Statement preferred alternative and administration of Mitchell Act hatchery funding. NMFS Consultation Number: NWR-2014-697.

NWMT. 2001. Northwest Marine Technologies CWT Detection Manual. Northwest Marine Technology. WA.

Papaspiliopoulos, O., G. O. Roberts, and M. Sköld. 2007. A General Framework for the Parametrization of Hierarchical Models. Statistical Science 22 (1): 59–73.

Paquet, P.J., T. Flagg, A. Appleby, J. Barr, L. Blankenship, D. Campton, M. Delarm, T. Evelyn, D. Fast, J. Gislason, P. Kline, D. Maynard, L. Mobrand, G. Nandor, P. Seidel, and S. Smith. 2011. Hatcheries, conservation, and sustainable fisheries-achieving multiple goals: Results of the Hatchery Science Review Group's Columbia River basin review. Fisheries, 36(11): 547–561. https://doi.org/10.1080/03632415.2011.

Parken, C.K., R.E. Bailey, and J.R. Irvine. 2003. Incorporating uncertainty into area under the curve and peak count salmon escapement estimation. North American Journal of Fisheries Management 23:78-90.

Parsons, A. L., and J. R. Skalski. 2009. Quantitative assessment of salmonid escapement techniques. Reviews in Fisheries Science, 18(4): 301-314.

Rawding, D., L. Brown, B. Glaser, S. VanderPloeg, S. Gray, C. Gleizes, P. Hanratty, J. Holowatz, and T. Buehrens. 2014. Coho Salmon Escapement Estimates and Coded-Wire-Tag Recoveries in Washington’s Lower Columbia River Tributaries in 2010. In D. Rawding, B. Glaser, & T. Buehrens (Eds.), Lower Columbia River fisheries and escapement evaluation in southwest Washington, 2010. FPT 14-10. Washington Department of Fish and Wildlife, Southwest Region.

Quinn, T. P. 2018. The Behavior and Ecology of Pacific Salmon and Trout. University of Washington Press, Vancouver, Canada. ISBN-10: 0295743336.

Rawding, D., S. VanderPloeg, A. Weiss, and D. Miller. 2010. Preliminary Spawning Distribution of Tule Fall Chinook Salmon in Washington's portion of the Lower Columbia River Evolutionary Significant Unit Based on Field Observation, GIS Attributes, and Logistic Regression. Washington Department of Fish and Wildlife. Olympia, WA. 17pp.

Riecke, T. V. P. J. Williams, T. L. Behnke, D. Gibson, A. G. Leach, B. S. Sedinger, P. A. Street., and J. S. Sedinger. 2019. Integrated population models: Model assumptions and inference. Methods in Ecology and Evolution, 10 (7): 1072-1082.

Scheuerell, M. D., E. R. Buhle, B. X. Semmens, M. J. Ford, T. Cooney, and R. W. Carmichael. 2015. Analyzing large-scale conservation interventions with Bayesian hierarchical models: a case study of supplementing threatened Pacific salmon. Ecology and Evolutuion, 5(10): 2115-2125.

Schindler, D. E., R. Hilborn, B. Chasco, C. P. Boatright, T. P. Quinn, L. A. Rogers, and M. Webster. 2010. Population diversity and the portfolio effect in an exploited species. Nature, 465: 609–612.

Schwarz, C.J., and G. G. Taylor. 1998. The use of the stratified-Petersen estimator in fisheries management: estimating pink salmon (Oncorhynchus gorbuscha) in the Frazier River. Can. J. Fish. Aquat. Sci. 55:281-297.

Stevens, D. L. 2002. Sampling design and statistical methods for the integrated biological and physical monitoring of Oregon streams. Oregon State University, Department of

Statistics, Corvallis, Oregon, and EPA National Health and Environmental Effects Research Laboratory, Western Ecology Division, Corvallis, Oregon.

Temple, G. M., and T. N. Pearsons. 2007. Electrofishing: Backpack and Drift Boat *in* D. H. Johnson, B. M. Shrier, J. S. O'Neal, J. A. Knutzen, X. Augerot, T. A. O-Neil, and T. N. Pearsons, editors. Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations. American Fisheries Society, Bethesda, Maryland.

Walther, E. J., M. S. Zimmerman, and P. A. H. Westley. 2021. Landscape attributes explain salmonid ecological neighbourhoods across a complex river network. Freshwater Biology, 66 (10): 1930-1948.

Wilson, J., and T. W. Buehrens. 2024. Weirs: An effective tool to reduce hatchery–wild interactions on the spawning grounds? North American Journal of Fisheries Management, 44: 21-38.

## 2.G. Acknowledgements

This work would not have been possible without the efforts of countless WDFW staff who walked tens of thousands of miles of streams counting fish and redds and handled countless fish at weirs and traps. Thanks to the Bonneville Power Administration, the Pacific Coastal Salmon Recovery Fund, Tacoma Power, PacificCorp, and NOAA Fisheries for funding this work. Thanks to Danny Warren for building the data systems behind the scenes that make this project work.

## 2.H. Tables

Table 1. Coho salmon populations comprising the Lower Columbia River ESU located within Washington State, subpopulation monitoring areas, and methods used between 2010 and 2018 to monitor these populations. Map indexes correspond to the number and letter codes used to identify populations, weirs, and trap-and-haul locations in Figure 1. Monitoring areas where a at least a portion of the spawning ground surveys were representatively identified using the GRTS algorithm have an asterisk (*) next to their name; spawning ground surveys in areas without an asterisk had complete coverage and therefore did not need GRTS surveys to facilitate extrapolating densities to un-surveyed areas.

| Population | | | | Weirs | | Trap-and-haul | | Spawning Ground Survey Years |
|---|---|---|---|---|---|---|---|---|
| Name | Map Index | Monitoring Areas | Km of habitat | Map Index | Years | Map Index | Years | |
| Grays, Chinook | 1 | Grays Chinook* | 262 | | | | | 2010-22 |
| Elochoman, Skamokawa | 2 | Elochoman Skamokawa* | 161 | | | | | 2010-22 |
| Mill, Abernathy, Germany, Coal | 3 | Abernathy | 37 | F | 2010-12 | | | 2010-22 |
| Mill, Abernathy, Germany, Coal | 3 | Upper Abernathy | 21 | E | 2013-19 | | | 2010-22 |
| Mill, Abernathy, Germany, Coal | 3 | Mill | 39 | | | | | 2010-22 |
| Mill, Abernathy, Germany, Coal | 3 | Germany | 20 | | | | | 2010-22 |
| Upper Cowlitz | 8 | Upper Cowlitz and Cispus | 219 | | | B | 2010-22 | |
| Cispus | 10 | Upper Cowlitz and Cispus | 117 | | | B | 2010-22 | |
| Tilton | 9 | Tilton | 200 | | | A | 2010-22 | |
| Lower Cowlitz | 4 | Lower Cowlitz* | 329 | | | | | 2010-11, 13-22 |
| Lower Cowlitz | 4 | Delameter above weir | 24 | C | 2013-19 | | | 2010-11, 13-22 |
| Lower Cowlitz | 4 | Lacamas above weir* | 103 | A | 2014-22 | | | 2010-11, 13-22 |
| Lower Cowlitz | 4 | Olequa above weir* | 167 | B | 2013-22 | | | 2010-11, 13-22 |
| Lower Cowlitz | 4 | Ostrander above weir | 19 | D | 2013-19 | | | 2010-11, 13-22 |
| NF Toutle | 6 | Green* | 77 | | | | | 2010-22 |
| NF Toutle | 6 | NF Toutle (upper) | 81 | | | C | 2010-22 | |
| NF Toutle | 6 | NF Toutle (lower)* | 71 | | | | | 2010-22 |
| SF Toutle | 7 | SF Toutle* | 91 | | | | | 2010-22 |

|  |  |  |  |  |  |  |  |  |
|---|---|---|---|---|---|---|---|---|
| Coweeman | 5 | Coweeman* | 96 |  |  |  |  | 2010-22 |
| Kalama | 11 | Kalama* | 52 |  |  |  |  | 2010-22 |
| NF Lewis | 12 | Lower Lewis* | 74 |  |  |  |  | 2010-22 |
| NF Lewis | 12 | NF Lewis above Swift | 135 |  |  | D | 2012-22 |  |
| NF Lewis | 12 | Cedar* | 68 |  |  |  |  | 2010-22 |
| EF Lewis | 13 | EF Lewis* | 141 |  |  |  |  | 2010-22 |
| Salmon Creek | 14 | Salmon Creek* | 282 |  |  |  |  | 2011-22 |
| Washougal | 15 | Washougal* | 89 |  |  |  |  | 2010-22 |
| Lower Gorge | 16 | Lower Gorge* | 50 |  |  |  |  | 2010-22 |
| Lower Gorge | 16 | Duncan above weir | 3 | G | 2010-22 |  |  | 2010-22 |
| Upper Gorge | 17 | (not monitored) | NA |  |  |  |  |  |

Table 2. Priors used in the model.

| Parameter | Prior Distribution | Definition |
|---|---|---|
| $\sigma_\eta$ | half-Cauchy (0, 0.5) | Regional-scale wild spawner process variation sd |
| $z_{\eta_y}$ | Unit Normal | Scaled regional-scale wild spawner annual process error |
| $\mu_\lambda$ | Inv Gamma (1, 0.125) | Random effect mean for the vector of subpopulation wild abundance process error sds |
| $\sigma_{\lambda_\sigma}$ | Unit Normal | Random effect sd for the vector of subpopulation wild abundance process error sds |
| $\boldsymbol{z_\sigma}$ | Unit Normal | A *p*-length vector of scaled random effects for wild abundance process error sds |
| $\boldsymbol{z}_\lambda$ | Unit Normal | A *y* by *p* matrix of scaled subpopulation wild abundance process errors |
| $\sigma_H$ | half-Cauchy (0, 0.5) | Global process error sd for proportion of hatchery spawners |
| $\mathbf{z}_H$ | Unit Normal | A *y* by *p* matrix of process errors for proportion of hatchery origin spawners |
| $\log\left(\boldsymbol{\lambda}_{y=1}\right)$ | $\mu_{\lambda 1} + \boldsymbol{z}_{\lambda 1}\sigma_{\lambda 1}$ | Wild spawner density in year 1 (hierarchical prior) |
| $\mu_{\lambda 1}$ | Normal (3, 1.5) | Hyper-mean for wild spawner density in year 1 (95% coverage of 1-351 redds per km) |
| $\sigma_{\lambda 1}$ | half-Cauchy (0, 0.5) | Hyper-sd for wild spawner density in year 1 |
| $\boldsymbol{z}_{\lambda 1}$ | Unit Normal | Scaled random effects for wild spawner density in year 1 |
| $\mathrm{logit}\left(\boldsymbol{p}_{H_{y=1}}\right)$ | Normal (0, 3) | Vector of length *p* of the logit proportion of hatchery spawners in year 1 |
| $\sigma_F$ | half-Cauchy (0, 0.5) | Proportion female random effect sd |
| $\mu_F$ | Beta (0.5, 0.5) | Proportion female random effect mean. |
| $\boldsymbol{z_F}$ | Unit Normal | A *y* by *p* matrix of scaled proportion female random effects |
| $\sigma_M$ | half-Cauchy (0, 0.5) | Above-dam fishing exploitation rate random effect sd |
| $\mu_{\boldsymbol{M}}$ | Beta (0.5, 0.5) | A *p*-length vector random effect medians for above dam exploitation rates |
| $\boldsymbol{z_M}$ | Unit Normal | A *y* by *s* matrix of scaled above-dam exploitation rate random effects where *s* is the number of subpopulations with above-dam fisheries (a subset of *p*) |
| $\gamma$ | Lognormal (0, 1) | The number of observable redds per female spawner |
| $p_M$ | Beta (0.5, 0.5) | The proportion of fish that were tagged while passing weirs |
| $\mu_D$ | Lognormal (0, 1) | The random effect mean negative binomial sd (transformed dispersion) |

|  |  |  |
|---|---|---|
| $\sigma_{D_\sigma}$ | Unit Normal | The random effect sd for negative binomial sd (transformed dispersion) |
| $\boldsymbol{z_D}$ | Unit Normal | The *p*-length vector of scaled random effects for negative binomial dispersion parameters |
| $p_I$ | Beta (0.5,0.5) | Proportion of redds in index reaches for specific subpopulations and years |

Table 3. Descriptive statistics (median and interannual range) from 2010-2022 by Lower Columbia River coho salmon population for natural and hatchery-origin spawner abundance, wild spawner density, and the proportion of hatchery origin spawners (pHOS). Wild spawner abundance is compared with population recovery goals and the number (out of a possible 13 years), and percentage of years abundance exceeded the recovery goal is reported. Summary statistics reported in this table were calculated from posterior medians. The asterisk (*) for indicates that the recovery goal applies to spawners in both Washington and Oregon but only spawners in Washington were estimated.

| | Wild Spawners | | | | | Wild Spawner Density (fish * km$^{-1}$) | | pHOS | | Hatchery-Origin Spawners | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Population | Recovery Goal | yrs > goal | % yrs > goal | median | range | median | range | median | range | median | range |
| Coweeman | 1200 | 13 | 100% | 3501 | (1238-6876) | 36.66 | (12.96-72.00) | 0.15 | (0.04-0.25) | 539 | (129-1309) |
| EF Lewis | 2000 | 7 | 53.8% | 2408 | (544-4561) | 17.10 | (3.86-32.39) | 0.09 | (0.06-0.37) | 249 | (96-881) |
| Elochoman, Skamokawa | 2400 | 1 | 7.7% | 858 | (328-3079) | 5.32 | (2.03-19.07) | 0.36 | (0.18-0.72) | 486 | (168-2163) |
| Grays, Chinook | 2400 | 1 | 7.7% | 627 | (301-2689) | 2.39 | (1.15-10.27) | 0.58 | (0.40-0.93) | 895 | (288-3911) |
| Kalama | 500 | 0 | 0% | 75 | (31-268) | 1.45 | (0.61-5.19) | 0.75 | (0.51-0.96) | 343 | (88-1608) |
| Lower Cowlitz | 3700 | 9 | 69.2% | 4365 | (2474-24544) | 6.87 | (3.90-38.65) | 0.11 | (0.05-0.22) | 768 | (214-1558) |
| Lower Gorge | 1900* | 0 | 0% | 631 | (340-1537) | 11.92 | (6.42-29.02) | 0.16 | (0.06-0.27) | 108 | (51-559) |
| Mill, Abernathy, Germany, Coal | 1800 | 1 | 7.7% | 918 | (514-2480) | 9.60 | (5.37-25.92) | 0.12 | (0.02-0.28) | 161 | (14-555) |
| NF Lewis | 500 | 13 | 100% | 3350 | (1007-7739) | 12.09 | (3.63-27.92) | 0.62 | (0.15-0.84) | 6025 | (279-9165) |
| NF Toutle | 1900 | 6 | 46.20% | 1818 | (830-4567) | 7.93 | (3.62-19.91) | 0.22 | (0.14-0.61) | 419 | (271-2893) |
| Salmon Creek | NA | NA | NA | 1952 | (953-4627) | 6.92 | (3.38-16.40) | 0.05 | (0.02-0.11) | 89 | (24-273) |
| SF Toutle | 1900 | 7 | 53.8% | 2036 | (1117-8364) | 22.32 | (12.24-91.71) | 0.11 | (0.06-0.49) | 255 | (91-2002) |
| Tilton | NA | NA | NA | 2401 | (899-8920) | 12.01 | (4.50-44.62) | 0.62 | (0.36-0.82) | 4397 | (1561-10895) |
| Upper Cowlitz and Cispus | 2000 | 8 | 61.5% | 3562 | (12-11532) | 10.62 | (0.04-34.37) | 0.67 | (0.52-1.00) | 8861 | (941-22620) |
| Washougal | 1500 | 0 | 0% | 423 | (151-1026) | 5.32 | (1.90-12.90) | 0.55 | (0.10-0.75) | 650 | (47-1787) |
| WA LCR ESU | 23700 | 10 | 76.9% | 28311 | (12348-85910) | 9.47 | (4.13-28.74) | 0.46 | (0.29-0.57) | 24397 | (9422-52863) |

## 2.I. Figures

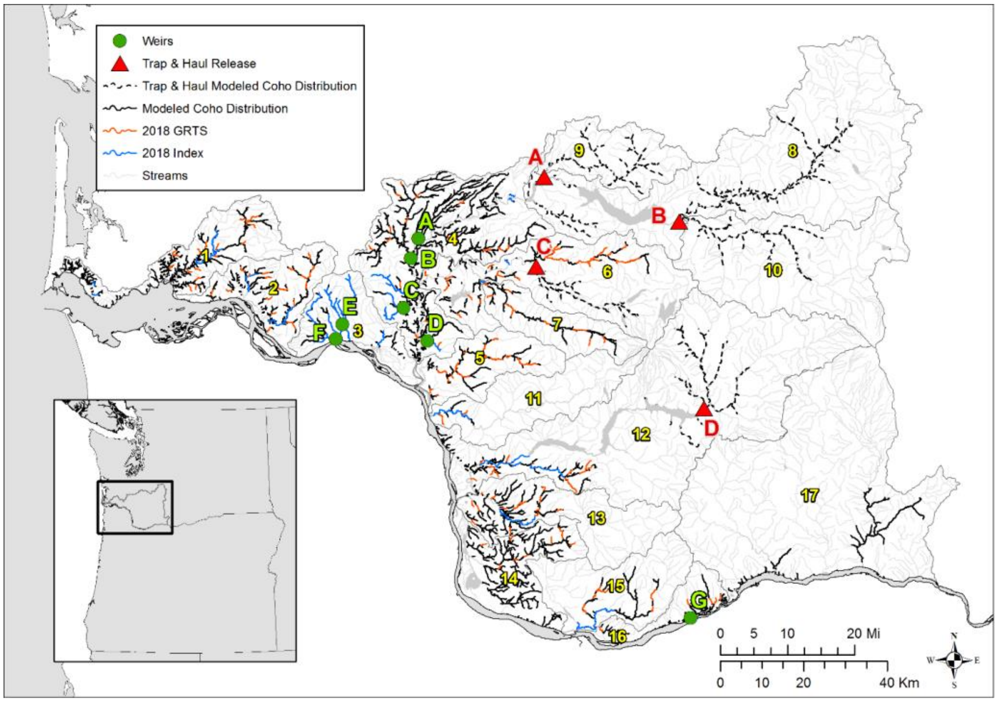


Figure 1. The portion of the Lower Columbia River Coho Salmon Evolutionarily Significant Unit within Washington State. Distinct populations are identified by gray borders and are identified by number. Black lines show the estimated extent of coho salmon occupancy in areas below (continuous) and above (dashed) dams where fish are passed upstream via trap-and-haul. Trap-and-haul release sites are denoted by triangles and partial capture weirs used for mark-recapture are denoted by circles. Spawning ground surveys reaches from 2018 are shown as an example (orange—representative GRTS reaches, blue—nonrandom index reaches). The mainstems of the Cowlitz River below Barrier Dam and North Fork Toutle River below the Sediment Retention Structure (a dam) were not monitored and are thought to experience low use by coho. The mainstem North Fork Lewis below Merwin Dam, and the Upper Gorge are used by coho but were not monitored in a manner that would facilitate abundance estimation at the population scale. See Table 1 for a key to number and letter codes.

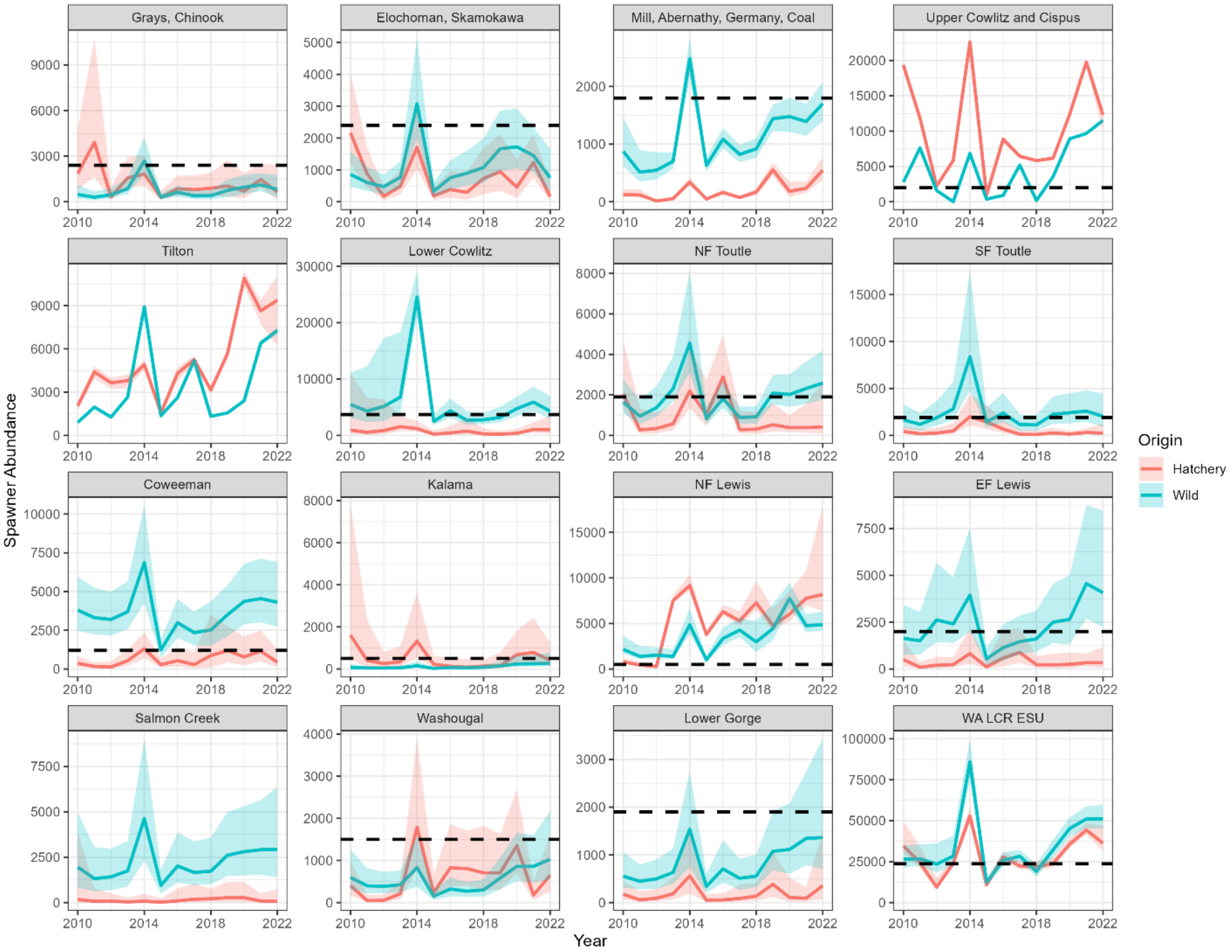


Figure 2. Abundance of wild and hatchery origin-coho salmon from 2010-2022 in each Washington population of the Lower Columbia River ESU. Bold colored lines are posterior medians and shading depicts 95% credible intervals. Bold dotted lines depict ESA recovery goals. Note that the recovery goal for the Lower Gorge population (and recovery goal) contains habitat in Oregon, which was not included in these estimates. Note variable scales on y-axis.

Figure 3. Spawner density (fish per km) of wild and hatchery-origin coho salmon from 2010-2022 in each Washington population of the Lower Columbia River ESU. Bold colored lines are posterior medians and shading depicts 95% credible intervals. Horizontal lines correspond to the number of coho salmon estimated to fully seed habitat based on a hockey stick recruitment model in Bradford (2000), converted from female spawners to total spawners using the lower Columbia mean proportion of females (0.47). Note that the Lower Gorge population contains habitat in Oregon, which was not included in these estimates. Note $\log_{10}$ scale on y-axis.

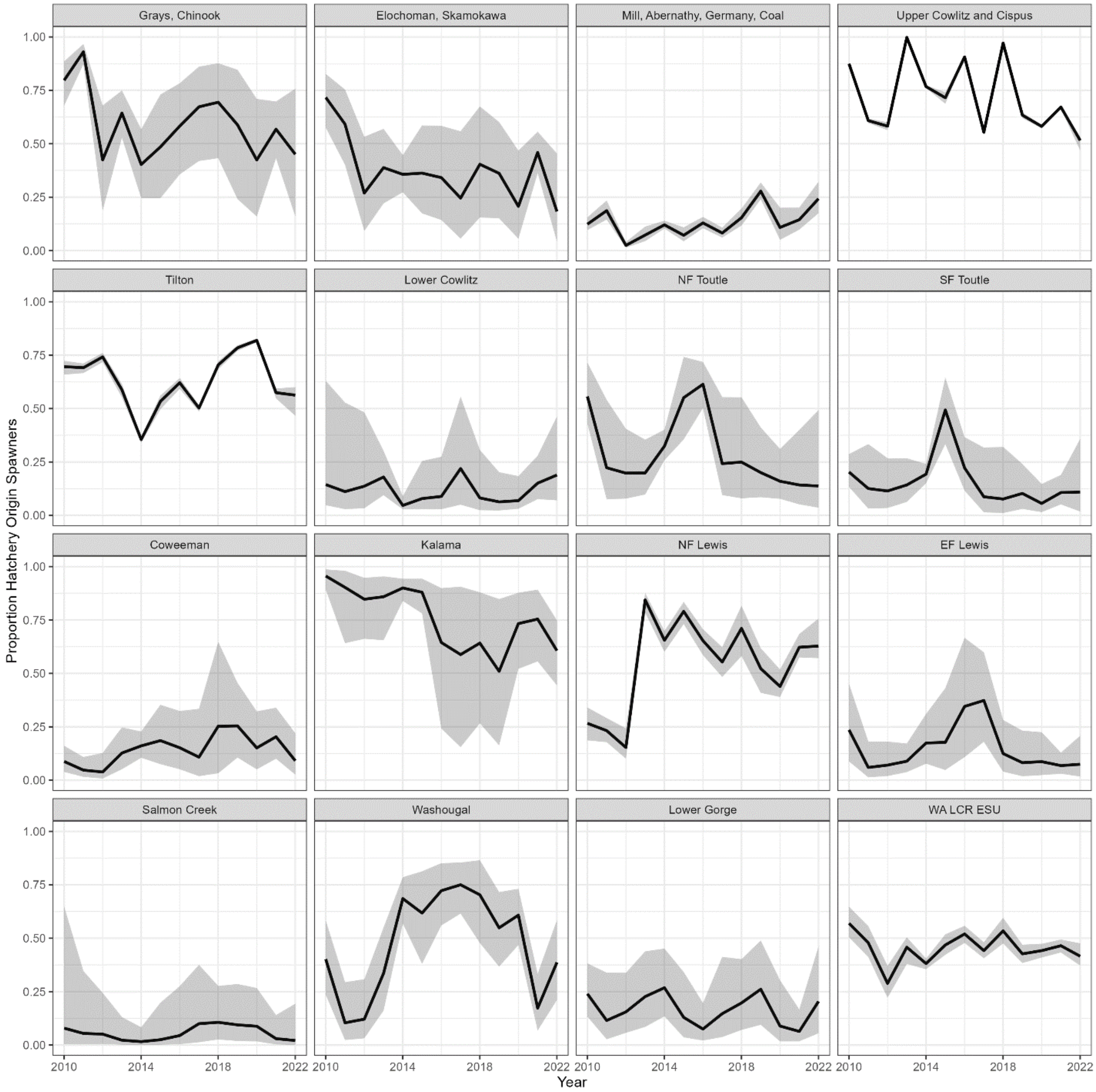


Figure 4. The proportion of hatchery-origin coho salmon from 2010-2022 within each Washington population of the Lower Columbia River ESU. Note that the Lower Gorge population contains habitat in Oregon, which was not included in these estimates. Bold colored lines are posterior medians and shading depicts 95% credible intervals.

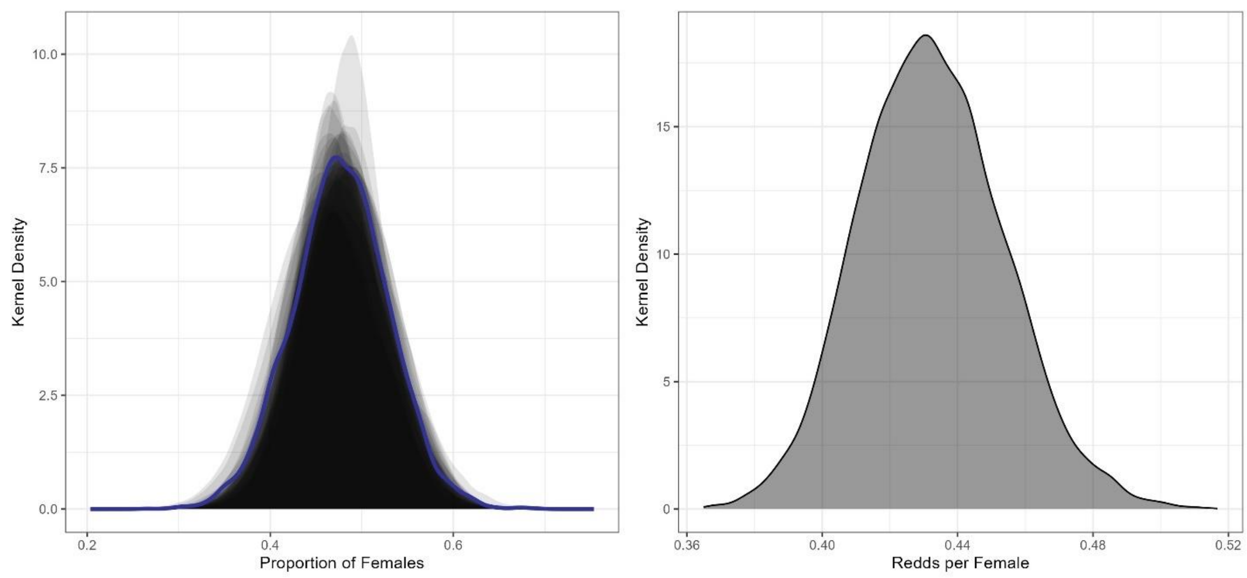


Figure 5. Posterior density plots for variables used to relate redd counts to spawner abundances. Year- and population-specific estimates of the proportion of females are shown (black shading) as well as the predictive distribution generated from the random effect hyper-parameters (blue line) in the left panel. The apparent (observable) number of redds per female, $\gamma$, is shown in the right panel.

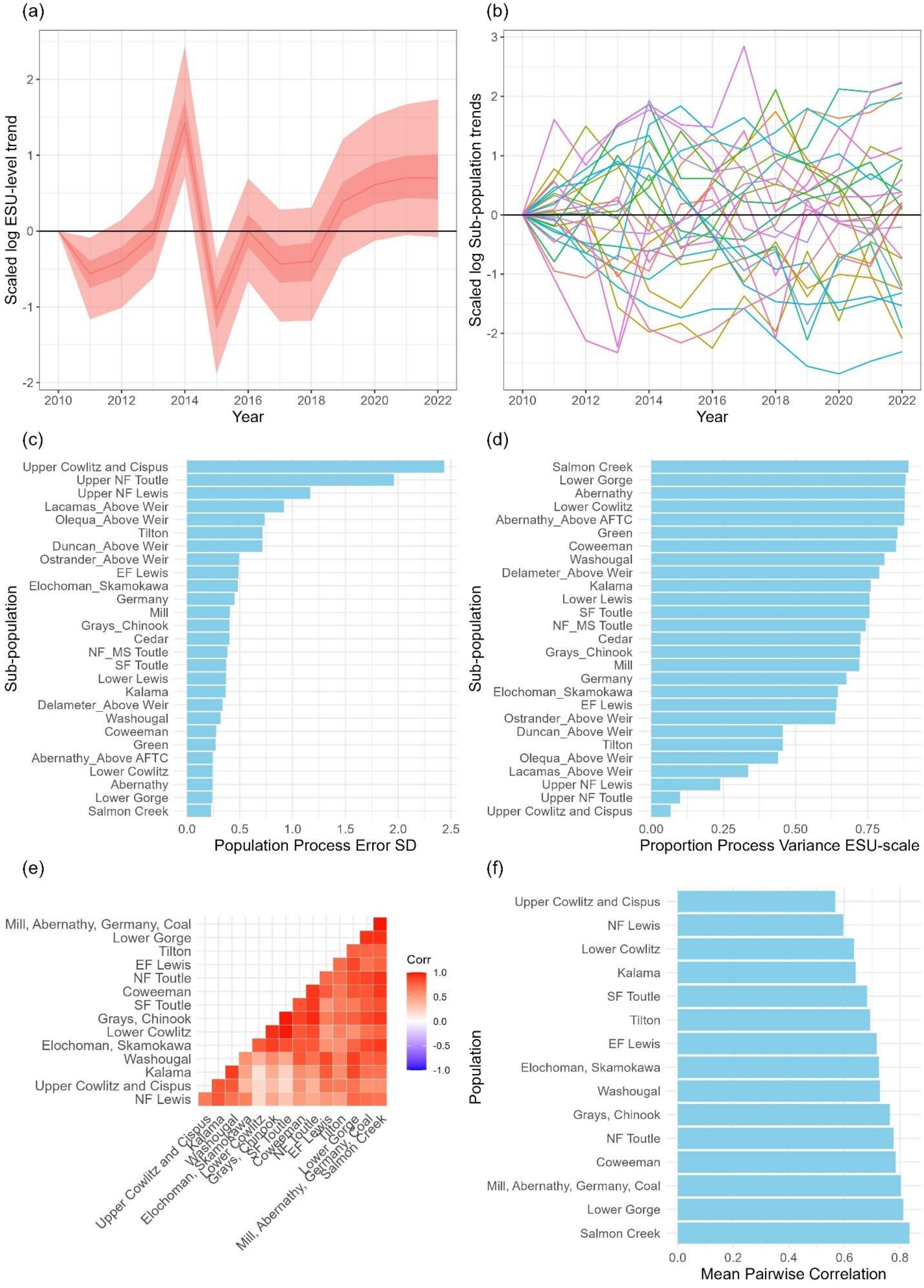

(a)
Scaled log ESU-level trend
Year
2010
2012
2014
2016
2018
2020
2022
2
1
0
-1
-2
(b)
Scaled log Sub-population trends
Year
3
2
1
0
-1
-2
2010
2012
2014
2016
2018
2020
2022
(c)
Sub-population
Population Process Error SD
Upper Cowlitz and Cispus
Upper NF Toutle
Upper NF Lewis
Lacamas_Above Weir
Olequa_Above Weir
Tilton
Duncan_Above Weir
Ostrander_Above Weir
EF Lewis
Elochoman_Skamokawa
Germany
Mill
Grays_Chinook
Cedar
NF_MS Toutle
SF Toutle
Lower Lewis
Kalama
Delameter_Above Weir
Washougal
Coweeman
Green
Abernathy_Above AFTC
Lower Cowlitz
Abernathy
Lower Gorge
Salmon Creek
0.0
0.5
1.0
1.5
2.0
2.5
(d)
Sub-population
Proportion Process Variance ESU-scale
Salmon Creek
Lower Gorge
Abernathy
Lower Cowlitz
Abernathy_Above AFTC
Green
Coweeman
Washougal
Delameter_Above Weir
Kalama
Lower Lewis
SF Toutle
NF_MS Toutle
Cedar
Grays_Chinook
Mill
Germany
Elochoman_Skamokawa
EF Lewis
Ostrander_Above Weir
Duncan_Above Weir
Tilton
Olequa_Above Weir
Lacamas_Above Weir
Upper NF Lewis
Upper NF Toutle
Upper Cowlitz and Cispus
0.00
0.25
0.50
0.75
(e)
Mill, Abernathy, Germany, Coal
Lower Gorge
Tilton
EF Lewis
NF Toutle
Coweeman
SF Toutle
Grays, Chinook
Lower Cowlitz
Elochoman, Skamokawa
Washougal
Kalama
Upper Cowlitz and Cispus
NF Lewis
Upper Cowlitz and Cispus
Kalama
Washougal
Elochoman, Skamokawa
Lower Cowlitz
Grays, Chinook
SF Toutle
Coweeman
NF Toutle
EF Lewis
Tilton
Lower Gorge
Mill, Abernathy, Germany, Coal
Salmon Creek
Corr
1.0
0.5
0.0
-0.5
-1.0
(f)
Population
Mean Pairwise Correlation
Upper Cowlitz and Cispus
NF Lewis
Lower Cowlitz
Kalama
SF Toutle
Tilton
EF Lewis
Elochoman, Skamokawa
Washougal
Grays, Chinook
NF Toutle
Coweeman
Mill, Abernathy, Germany, Coal
Lower Gorge
Salmon Creek
0.0
0.2
0.4
0.6
0.8

Figure 6. The contribution of local (sub-population and population) and regional trends to interannual variation in wild spawner abundance. The ESU-level (a) and sub-population-level (b) scaled log trends in wild spawner density were calculated as the cumulative sum of the ESU-level and sub-population-level wild spawner density process errors, respectively. The sub-population level process error standard deviations (c) identify the scale of variation in the sub-population component of abundance trends, which expressed a as a proportion of total process variation (d) differed considerably among populations. The magnitude of sub-population variation relative to ESU-level variation ultimately determined how correlated abundance time series were among populations (e-f).

## 2.J. Supplementary Materials

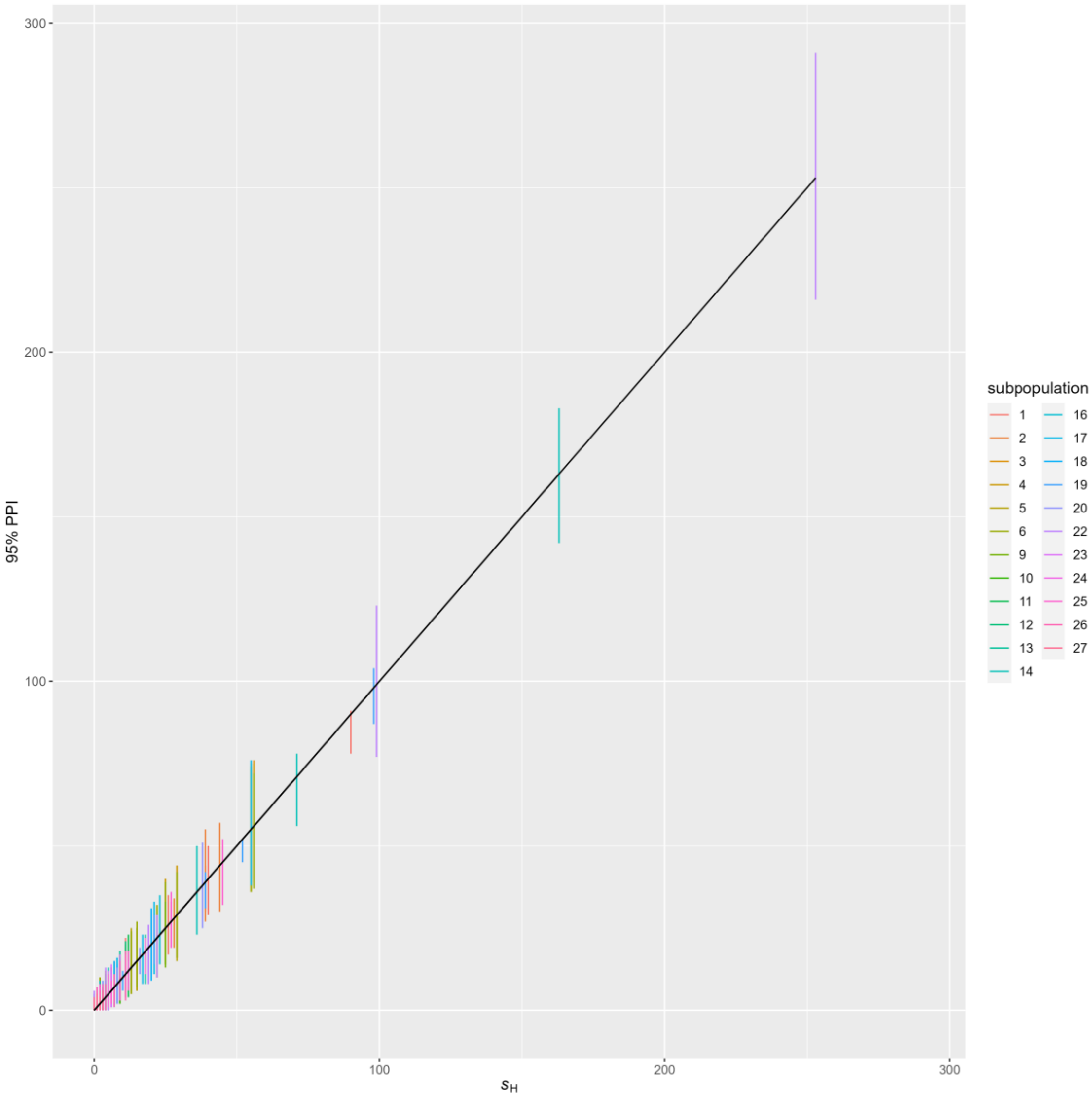


Figure S1. Regression of observed vs. 95% posterior predictive interval () for hatchery origin spawner count $s_H$. Black lines are the 1:1 line. Lack of fit is indicated by PPI that do not overlap the black line.

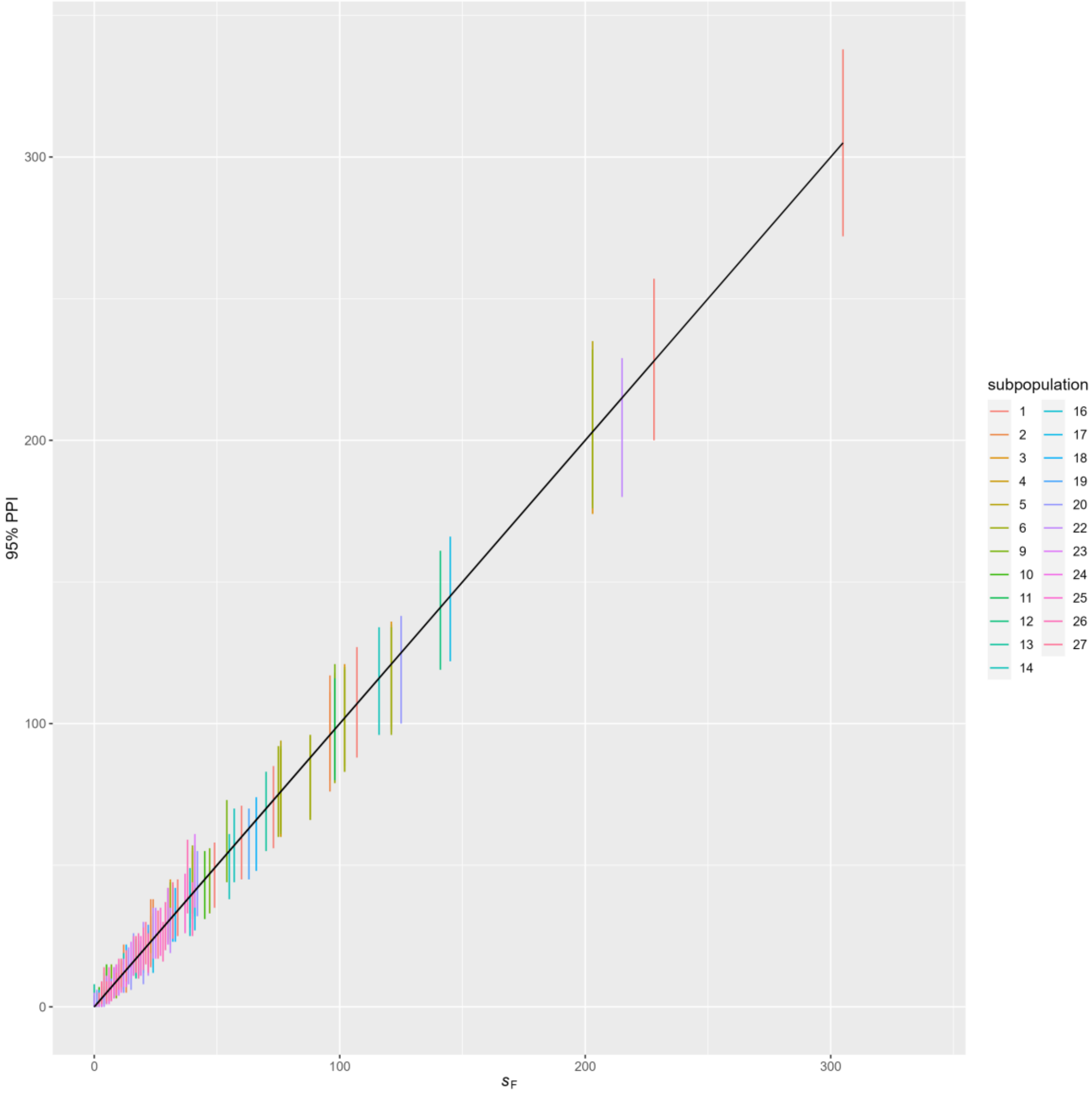


Figure S2. Regression of observed vs. 95% posterior predictive interval for female spawner count $s_F$. Lack of fit is indicated by PPI that do not overlap the black line.

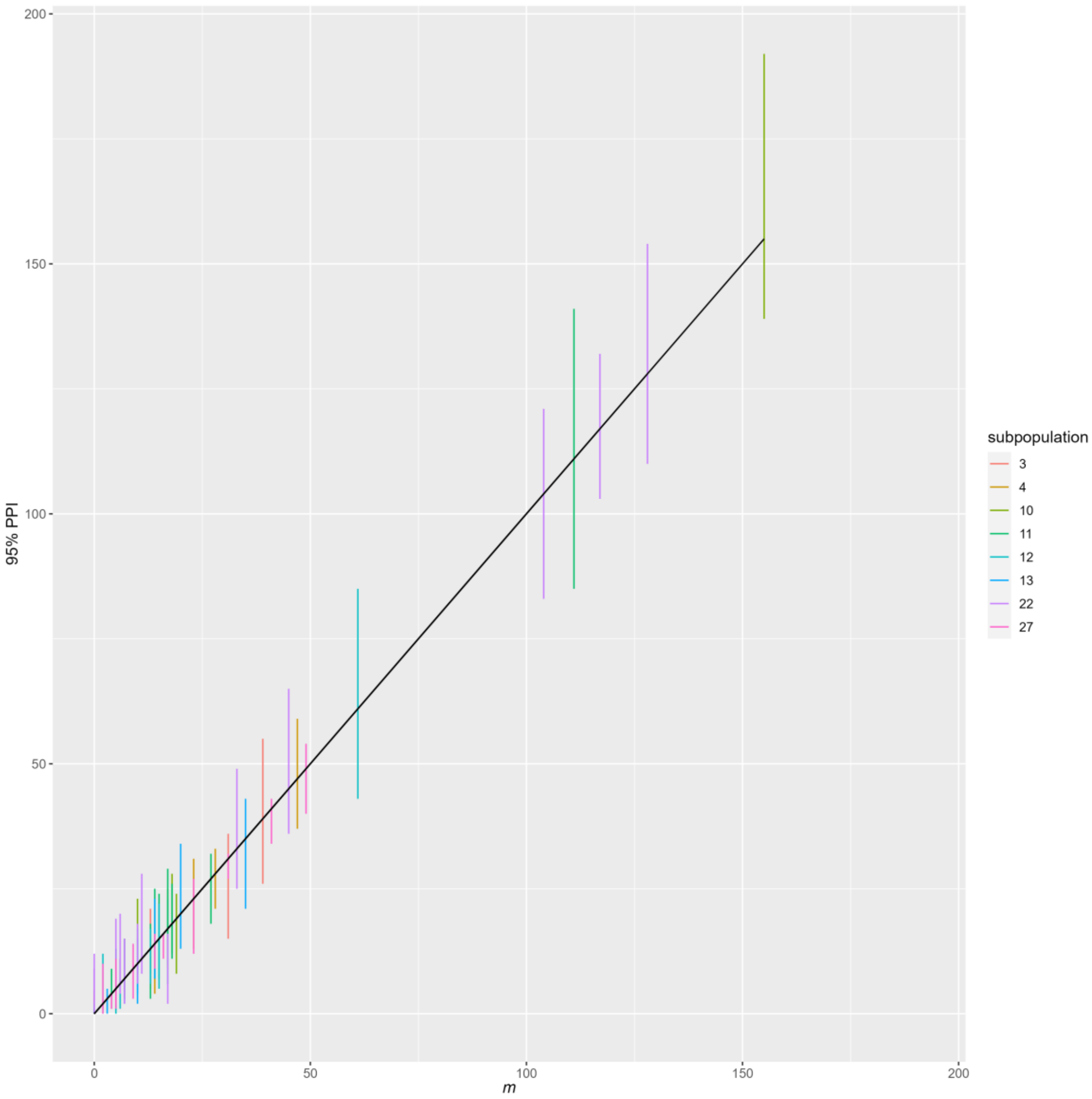


Figure S3. Regression of observed vs. 95% posterior predictive interval for recaptures $m$ used in mark recapture estimates of spawner abundance. Lack of fit is indicated by PPI that do not overlap the black line.

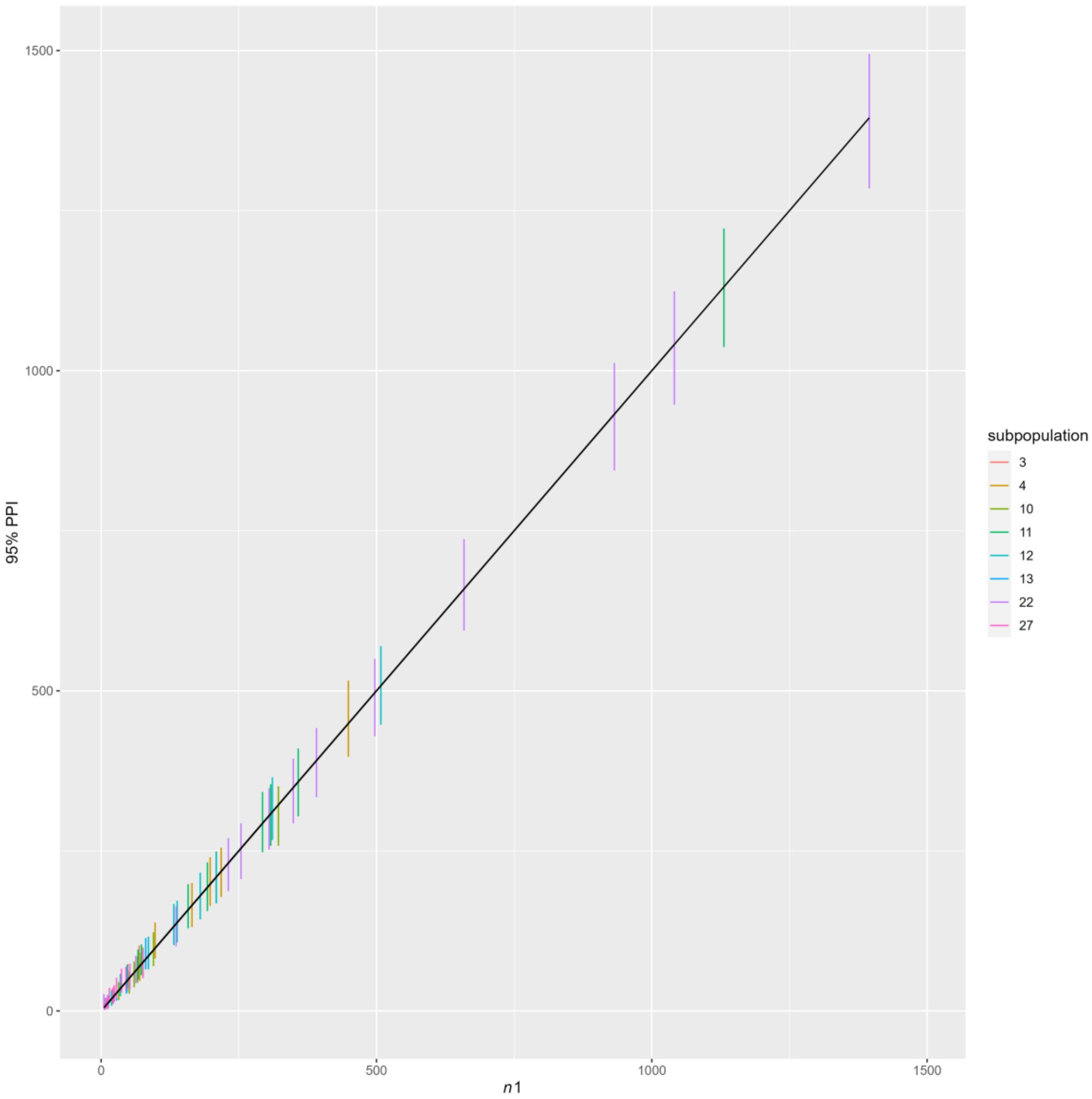


Figure S4. Regression of observed vs. 95% posterior predictive interval for for marks $n1$ used in mark recapture estimates of spawner abundance. Lack of fit is indicated by PPI that do not overlap the black line.

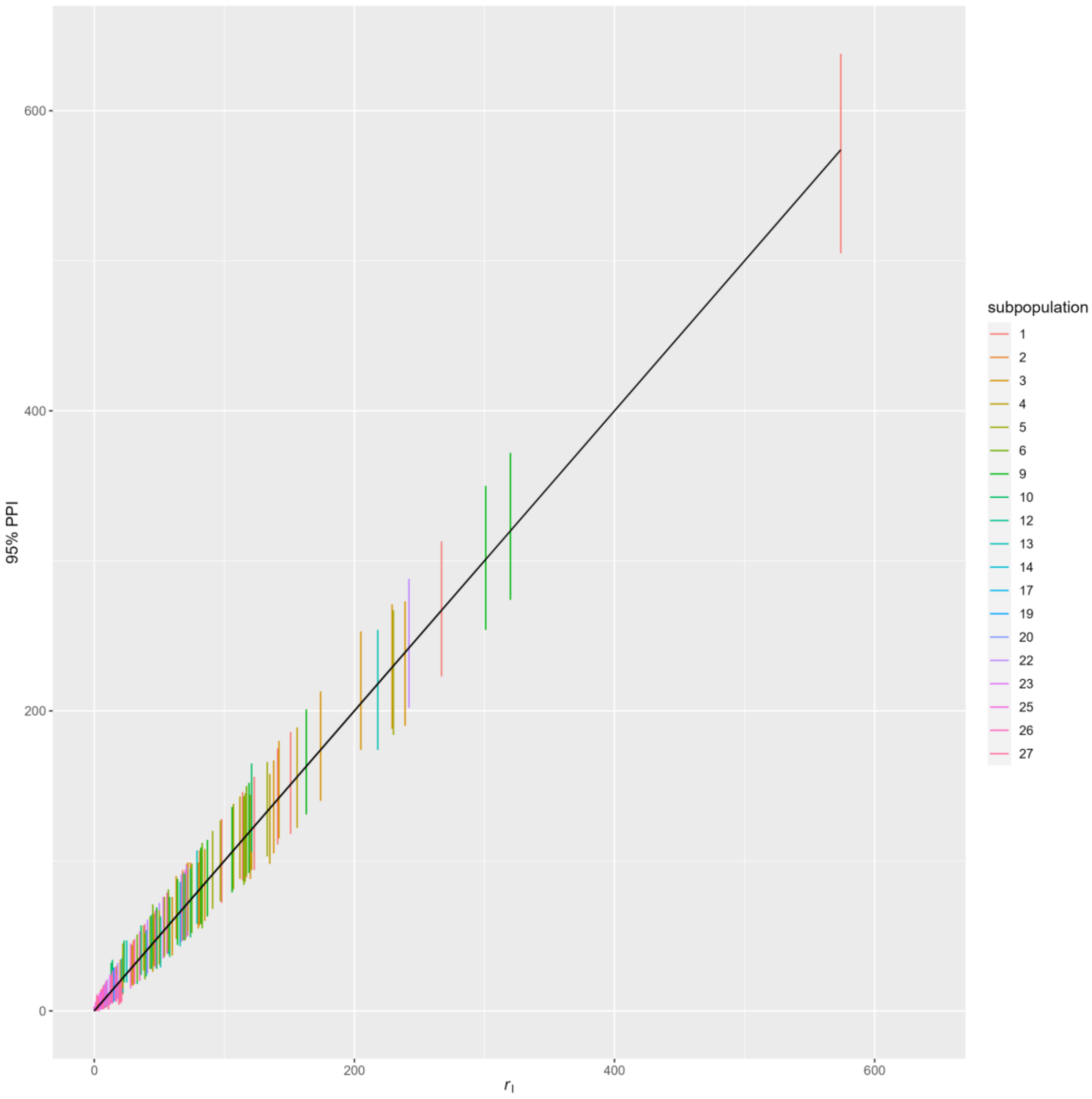


Figure S5. Regression of observed vs. 95% posterior predictive interval for index redd counts $r_I$. Lack of fit is indicated by PPI that do not overlap the black line.

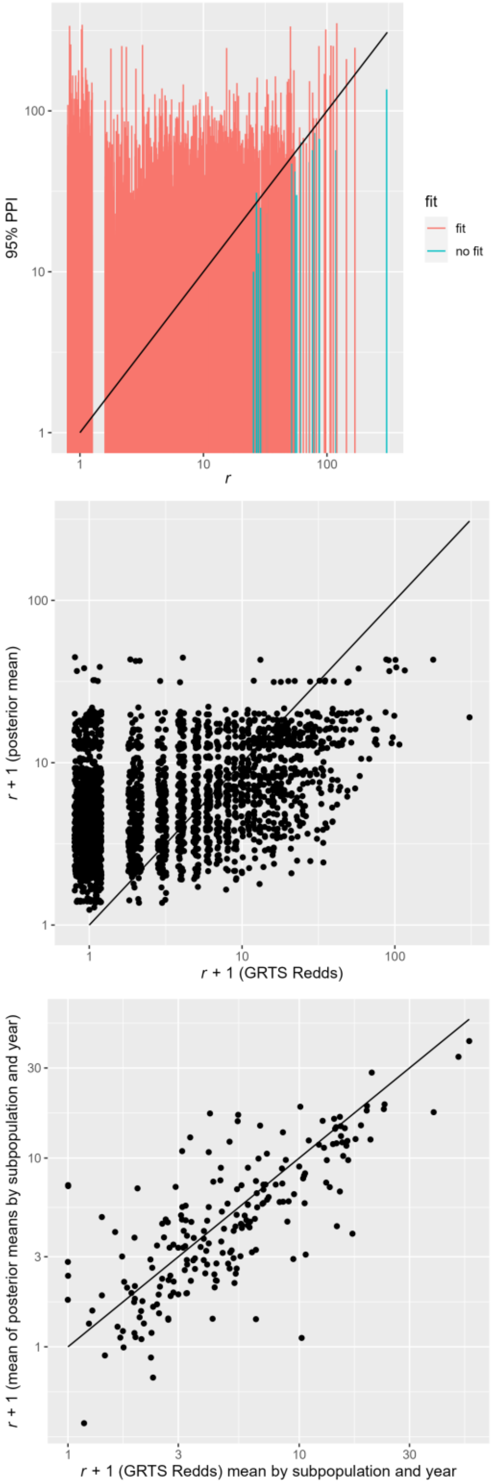


Figure S6. Regression of observed vs. posterior predictive distribution of GRTS redd counts $r$. Both x and y axes are log10 transformed and a constant of 1 was added to observations and predictions to aid visualization. A 1:1 line is shown in black. Top panel: x values are jittered observations (redd count in an individual subpopulation, year, and reach) and y are 95% PPI. Coloring depicts observations within and outside of the 95% PPI. Middle Panel: x values are jittered observations (redd count in an individual subpopulation, year, and reach) and y are the mean of the posterior predictive distribution. Bottom Panel: x values are the mean of redd counts across all reaches in an individual subpopulation and year and y are mean across reaches of the posterior predictive redd total by subpopulation and year. Overall, 2786 out of 2828 observations (98.5%) were within the 95% PPI. When this calculation is performed separately for each year and subpopulation, 181 of 207 year-by-subpopulation combinations had >=95% of observations within the 95% PPI. The remainder had less, and of these the lowest coverage was 75% of year and subpopulation specific reach level redd observations were within the 95%PPI. Although the top panel shows good PPI interval coverage, it appears to show lack of correspondence between the observed and expected. This is a natural effect of a shrinkage estimator (multivariate state space process model) applied to a high noise low mean count process. The middle plot shows that at the observation level, the posterior means were associated with the observations. The bottom panel shows that at the subpopulation and year level, the mean of the observations and the mean of the predictions were not different from 1:1. This suggests that any appearance of lack of association in the top panel was not pathological, but merely the model attempting to extract mean redd densities from very noisy low mean count data.

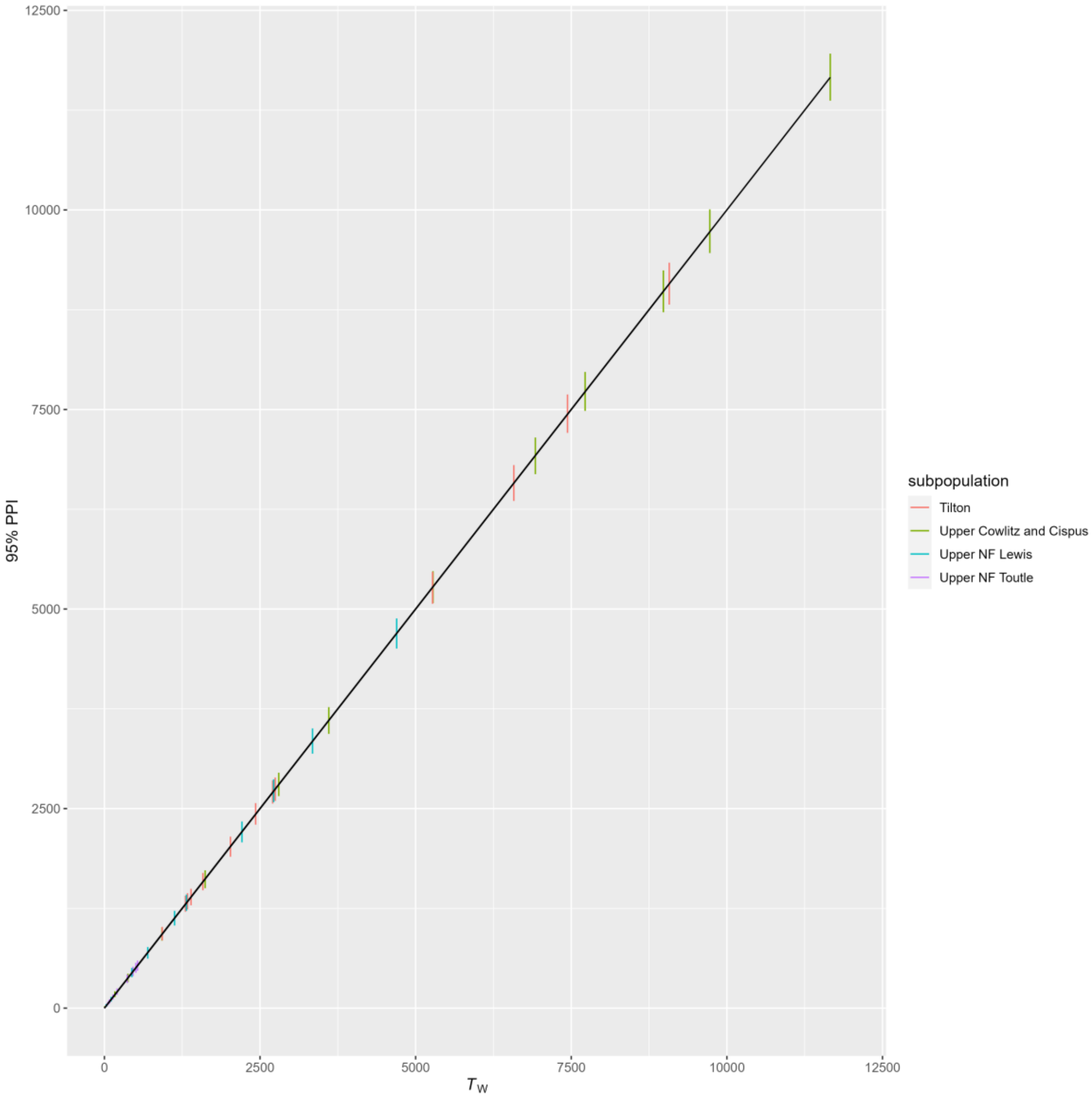


Figure S7. Regression of observed vs. 95% posterior predictive interval for the trap and haul census wild adult coho salmon count $T_W$. Lack of fit is indicated by PPI that do not overlap the black line.

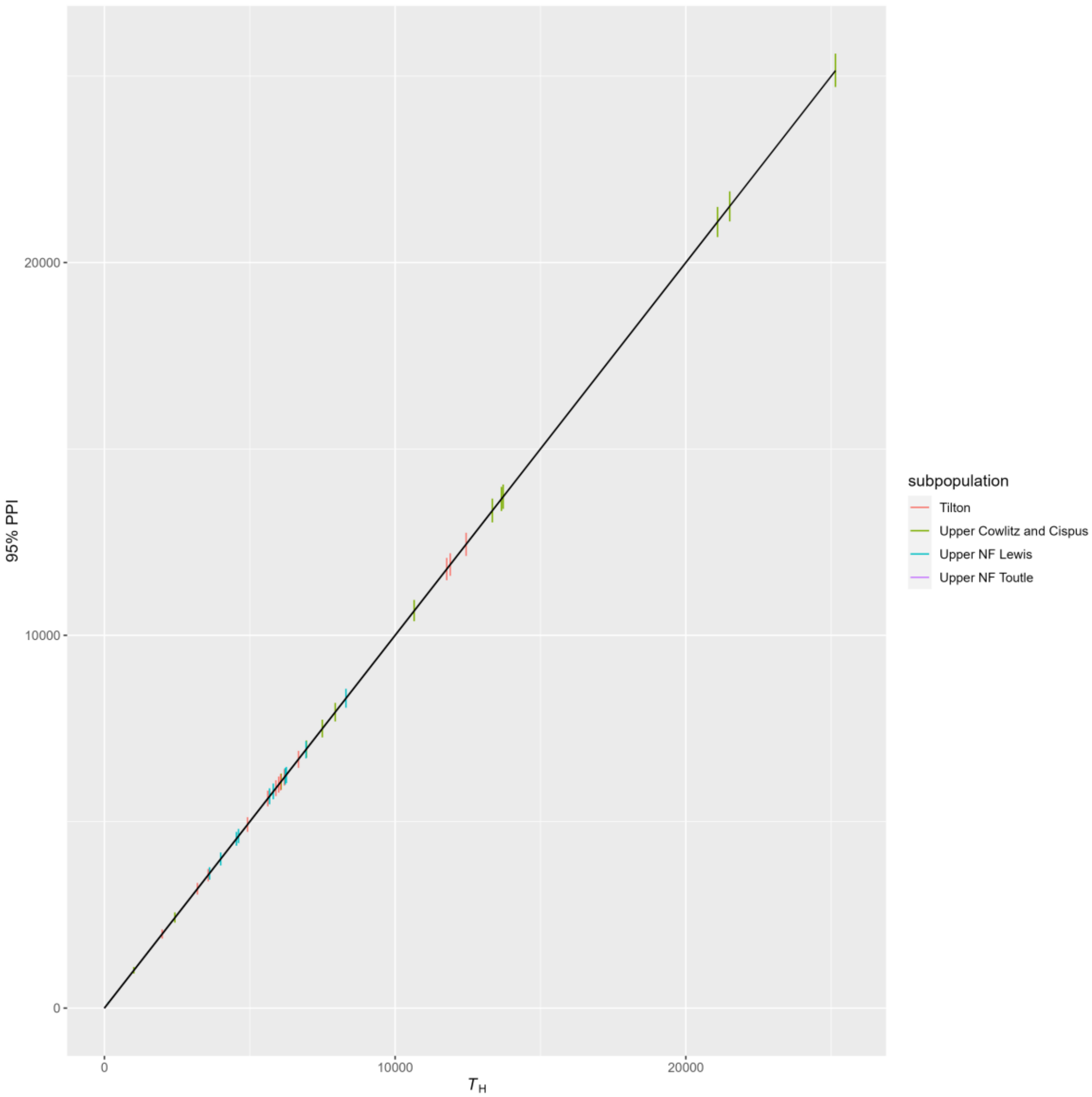


Figure S8. Regression of observed vs. 95% posterior predictive interval for the trap and haul census hatchery adult coho salmon count $T_H$. Lack of fit is indicated by PPI that do not overlap the black line.

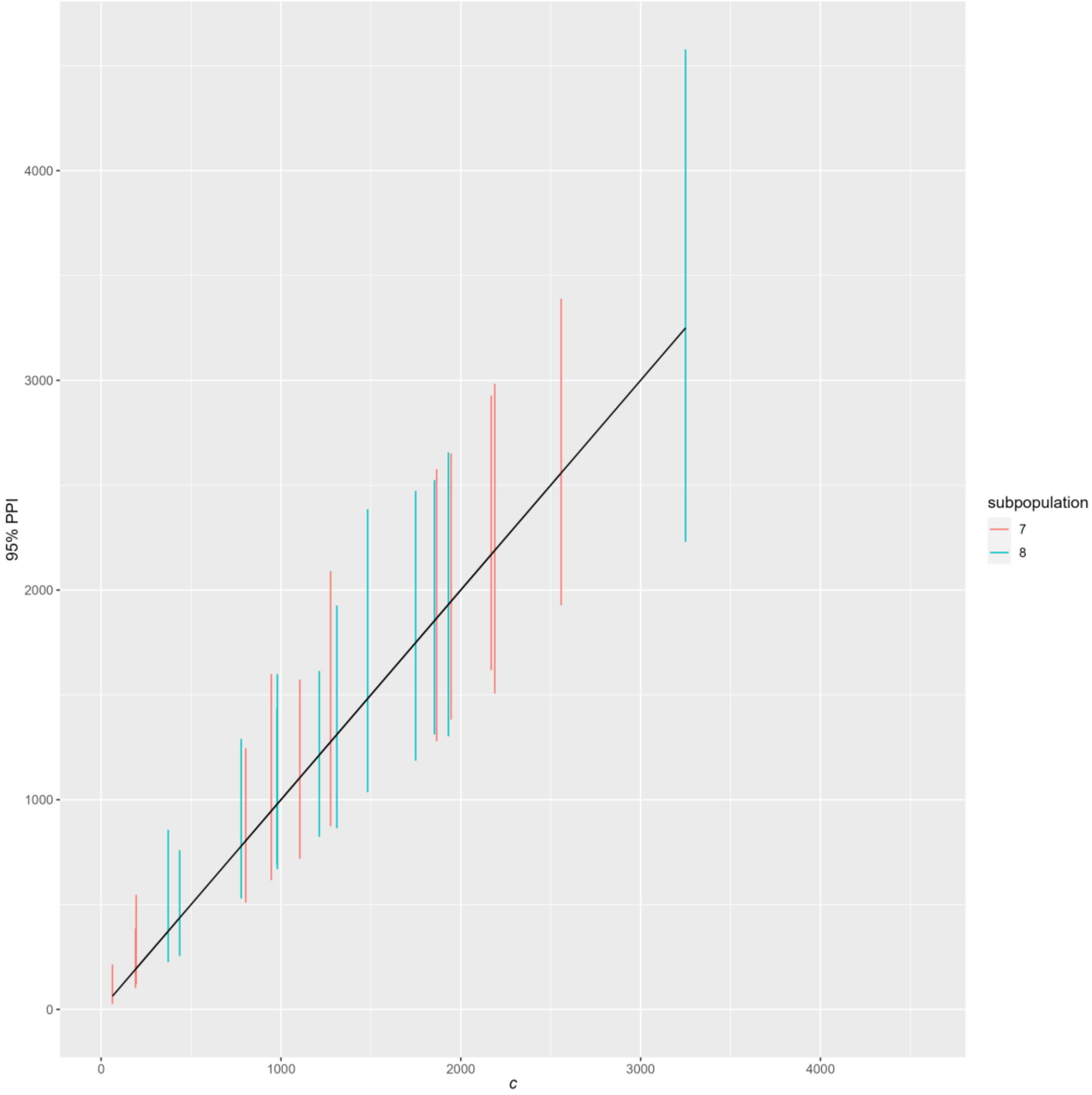


Figure S9. Regression of observed vs. 95% posterior predictive interval for the catch record card point estimate of hatchery adult coho salmon catch *c* used in to estimate exploitation rates occuring upstream of dams after the release of trap-and-haul adults. Lack of fit is indicated by PPI that do not overlap the black line.

# Chapter 3: A flexible model to estimate angler effort and catch with creel survey data in a multivariate state-space framework

## 3.A. Abstract

Sustainable management of exploited species requires quantification of mortality resulting from harvest. This is particularly true for imperiled species where subtle changes in exploitation can affect population viability. Creel surveys are a common method used to quantify recreational fishery catch by counting anglers to estimate fishing effort and interviewing anglers to estimate catch-per-unit-effort (CPUE). However, many statistical estimators commonly used to analyze creel survey data are inflexible and involve unreasonable assumptions, resulting in estimates of catch with poorly understood accuracy and precision. Sustainable management of recreational fisheries for Endangered Species Act (ESA) listed salmon (*Oncorhynchus* spp.) and steelhead trout (*O. mykiss*) necessitates the development of robust methods to quantify catch in an unbiased manner and quantify uncertainty. We developed a Bayesian multivariate state-space model to analyze creel survey data that flexibly accommodates multiple study designs and robustly quantifies uncertainty. We illustrate application of the model to data collected during the 2021 Skagit River wild winter steelhead catch-and-release sport fishery in Washington State, USA. We further apply model estimates of catch to a probabilistic management framework that quantifies opposing risks of overfishing and unnecessarily restricting fishing opportunity, showcasing how robust quantification of uncertainty can help fishery managers to balance these risks. Finally, we demonstrate how estimates of uncertainty enable optimizing study designs to balance desired precision with cost efficiency.

## 3.B. Introduction

Quantification of mortality resulting from harvest is essential to the sustainable management of exploited species (Hilborn and Walters 1992). Without estimates of exploitation, it is difficult to distinguish the population-dynamic effects of harvest from those of other causes. This is particularly important if the species under exploitation are depressed or already at risk for other reasons, such as low productivity (Stobutzki et al., 2001).

Creel surveys are a common method used to quantify harvest in recreational (sport) fisheries, thereby enabling sustainable management. Many creel survey methods were largely developed several decades ago (e.g., Malvestuto et al., 1978, Pollock et al. 1994) and typically estimate catch as the product of catch-per-unit-effort (CPUE) and fishing effort, both of which are observed during field surveys. These surveys can either occur at fixed access points where all anglers are presumed to pass during ingress and egress from the fishery (“access point” creels) or via creel samplers moving throughout the fishery area quantifying effort and interviewing anglers (“roving-roving” creels). The estimators developed for these study designs mostly use method-of-moments (MOM) and require stringent assumptions to be unbiased (e.g., asymptotic sample sizes and central limit theorem). The adequacy of such approaches is further worsened by frequent *ad hoc* modification of estimators by resource managers to cope with study designs that are unable to census fishing effort or because of data that fail to strictly conform to other study designs assumptions of traditional estimators. As a result, the accuracy and precision of harvest estimates derived from these approaches are often unknown.

Creel surveys can be used to achieve many objectives but often their primary purpose is to estimate recreational fishing effort and catch per unit effort (CPUE) that can subsequently be used to estimate catch (i.e., the number of fish captured and retained or released). Therefore, the goal of a creel survey is to collect data that characterizes the true variation in angler effort and catch rates over a specific period, which are influenced by a suite of underlying processes (e.g., abundance of fish, water temperature, turbidity, day of year, time of day, etc.). However, the true fluctuations in effort and catch rates can be obscured by variation in the data that is due to the sampling process. This uncertainty in measurements, or rather "inaccuracy" in observations, is referred to as observation error (De Valpine 2003).

Observation error occurs in many forms but typically arises in creel survey data for three reasons. First, while true fishing effort and catch rates are continuous variables, it is impossible to count fractions of anglers and for anglers to catch fractions of fish. Therefore, creel survey data will contain sampling error due solely to the fact that observations are discrete realizations of continuous processes. Second, creel survey data typically provide a snapshot of fishing effort at random times of the day and a sample of catch rates from a subset of anglers, and these random samples may imprecisely observe the mean daily effort and CPUE. Thus, creel estimates will almost always contain sampling error because fishing effort typically varies throughout the day and individual anglers will almost always have varying rates of success. Last, creel data can be collected imperfectly whereby anglers can be miscounted and anglers can misreport their catch and time spent fishing. Ultimately, the influence of each of these sources of error has on the estimates will depend on many factors including study design, characteristics of the fishery, and staff training. Nonetheless, if left unaccounted for, observation error can lead to noisy or biased estimates. Fortunately, analytical techniques have been developed that attempt to separate observation error and real "biological" signal, and one such approach is referred to as state-space modeling.

State-space models offer a robust and flexible framework for characterizing the dynamics of a system that is subject to observation error (De Valpine 2002). Specifically, a state-space model works by partitioning parameter estimation variance into two portions: the process model and the observation model. First, the process model provides a mathematical description of the factors that influence the true unobserved (i.e., latent) "states" (e.g., true daily catch rates), how the states are linked in time, and a way to account for fluctuations in the latent states that cannot be ascribed to measured covariates (i.e., process error). In contrast, the observation model specifies the relationships between the observed data and the unobserved states and aims to account for the error associated with the observation process. State-spaces models can be fit using various methods; however, Bayesian methodologies have become a popular choice due to their ability to characterize complex nonlinear models and their clear depiction of parameter uncertainty (De Valpine 2003; Clark and Bjørnstad 2004; Patterson et al. 2008).

There has recently been renewed interest in creel methods, with some authors applying the state-space framework to creel surveys. Several authors have adapted traditional creel methods such as roving-roving creel surveys to statistical models rather than MOM estimators (e.g., McCormick and Heckel 2023, Tucker et al. 2024, Trudeau et al., 2021). These models better emulate the generative processes underlying creel data, and thus generate more robust estimates of effort, catch, and uncertainty. However, Tucker et al. (2024) and Trudeau et al. (2021) focused on methods applicable to lakes, and McCormick and Heckel (2023) limited their work to quantification of effort rather than catch. As a result, there remain limited approaches to estimate catch using many of the study designs commonly employed by managers in riverine salmon and steelhead fisheries with modern statistical methods.

We developed a Bayesian state-space (BSS) model to analyze creel survey data with the goal of providing unbiased estimates of recreational catch that accurately quantifies uncertainty. Our BSS model allows for serial autocorrelation in fishing effort and catch rates among days and covariance in fishing effort and catch rates among angler types (e.g., bank, boat) and sections of a river. The current version of the model is designed to analyze creel survey data that were collected using a roving creel study design (Pollock et al. 1994). However, this BSS model could easily be reparametrized to accommodate other creel study designs (e.g., access survey). For illustration, we applied our BSS model to one season of creel survey data that were collected from a fishery in the Skagit River watershed located in northern Puget Sound, Washington State that principally involves catch-and-release of wild winter steelhead trout (*Oncorhynchus mykiss*). This fishery is of great public interest as it not only provides one of the few recreational fisheries for wild winter steelhead in the Puget Sound region but is also part of a larger population that has been listed as Threatened under the federal Endangered Species Act (ESA) since 2007. Finally, to illustrate the utility of the robust quantification of uncertainty enabled by our modeling framework, we applied probabilistic estimates of catch and effort from our model to two applied management applications: 1) we combined estimates of catch with preseason forecasts of run size to illustrate how managers can use estimates of uncertainty to balance opposing risks involved in fisheries management (i.e., overfishing vs. unnecessarily precluded fishing opportunity), 2) we re-sampled our data and re-fit the model illustrating how point estimates and estimates of uncertainty changed, illustrating how managers could use resampling to “tune” study designs for desired precision in order to optimize cost-efficiency.

## 3.C. Methods

### *Process Model*

We developed a model to emulate the theorized generative processes behind sport fishery catch. We assumed that catch arose as the product of two latent random variables: angling effort and catch-per-unit-effort (CPUE) as previous authors have assumed (e.g., Malvestuto et al., 1978, Pollock et al. 1994). These random variables were assumed to vary continuously in time and space in an autocorrelated manner, and that the product of their average values integrated over spatiotemporal strata determines the total catch in those strata. Given that both variables are zero-bounded, log-normal AR1 processes might provide a reasonable approximation of such processes. A minimally sufficient model for daily catch might simply include lognormal AR1 processes for the effort (e.g., numbers of anglers) and CPUE (e.g., catch per hour per angler) integrated over hours of fishing each day (e.g., hours per day). However, such a process ignores known sources of variability in angling effort and CPUE, such as between distinct spatial areas within a fishery, and between anglers employing different methods. Stratification is not necessary if observations of effort and CPUE could be sampled proportionally within each such stratum (i.e., *iid* sampling). However, CPUE estimates would be biased if more measurements of CPUE were inadvertently made in areas of higher CPUE or of “angler types” (hereafter “gear types”) with higher CPUE. Therefore, we developed a model that stratified effort and CPUE by spatial sections and by the gear type distinction that we thought was most likely to result in heterogenous CPUE and sampling rates; whether anglers were fishing from the bank or from boats.

The first unobserved state in the model was the expected hourly catch rate $\lambda^{C}$ was a gear type $G \times$ section $S$ -length state vector that was calculated for each day $d$. The logarithm of $\boldsymbol{\lambda_d^C}$ was modeled as the sum of the global mean of the log CPUE $\mu^{C}$ across all spatial sections and gear types, a $G \times S$ vector of unit-normal random effects $\boldsymbol{\gamma^C}$**,** which accounted for temporal variation in the mean of log CPUE among sections and gear types, and was rescaled by the among-strata standard deviation in CPUE $\sigma^{\gamma C}$, plus a $G \times S$ vector of daily CPUE residual $\boldsymbol{\omega^C}$ accounting for day-to-day deviations from the seasonal log means in each stratum (Table 1). This exponentiated sum was then multiplied by a section- and day-specific indicator variable $O_{d,s}$ identifying whether the fishery was open on a particular date. It was set to 1 for open dates and to 1E-09 for closed dates:

$$\boldsymbol{\lambda_d^C} = e^{\mu^{C} + \boldsymbol{\gamma^C} \sigma^{\gamma C} + \boldsymbol{\omega_d^C}} (o_{d,s}) \tag{1}$$

The daily catch rate residual $\boldsymbol{\omega_d^C}$ was assumed to be serially auto-correlated and thus modeled as a function of the residuals in the previous day multiplied by an auto-regressive (AR), mean-reverting lag-1 coefficient $\varphi^{C}$ plus a $G \times S$ process error vector $\boldsymbol{\varepsilon_d^C}$:

$$\boldsymbol{\omega_d^C} = \varphi^{C} \boldsymbol{\omega_{d-1}^C} + \boldsymbol{\varepsilon_d^C} \tag{2}$$

where the CPUE process errors were multivariate (across gears and sections) normally distributed, with mean zero and covariance matrix $\boldsymbol{\Sigma^C}$:

$$\boldsymbol{\varepsilon_d^C} \sim MVN(0, \boldsymbol{\Sigma^C}) \tag{3}$$

These process errors were modeled using a non-centered parameterization of the multivariate normal via decomposition of the covariance matrix:

$$\boldsymbol{\varepsilon}_{\boldsymbol{d}}^{\boldsymbol{C}} = \boldsymbol{L}^{\boldsymbol{C}}\boldsymbol{I}\sigma^{C}\boldsymbol{z}_{\boldsymbol{d}}^{\boldsymbol{C}} \tag{4}$$

where $\boldsymbol{L}^{\boldsymbol{C}}$ was the lower left triangular Cholesky factor of a square correlation matrix with a row dimension $G \times S$ and was multiplied by an identity matrix $\boldsymbol{I}$ of the same dimensions, which was multiplied by the daily CPUE process error standard deviation $\sigma^{C}$ to generate the Cholesky factor of $\boldsymbol{\Sigma}^{\boldsymbol{C}}$, which was multiplied by a $G \times S$ vector of unit normal daily process errors $\boldsymbol{z}_{\boldsymbol{d}}^{\boldsymbol{C}}$.

The second unobserved state, the expected effort (anglers) for each day $d$ was a gear types $G$ × sections $S$ state vector $\boldsymbol{\lambda}_{\boldsymbol{d}}^{\boldsymbol{E}}$. It was modeled as a log-linear sum of the seasonal global mean of the log effort $\mu^{E}$ across all spatial sections and gear types of the fishery and a $G \times S$-length vector of unit-normal random effects $\boldsymbol{\gamma}^{\boldsymbol{E}}$**,** which accounted for seasonal differences in the mean of log effort among sections and gear types, and was rescaled by the among strata-variance in CPUE $\sigma^{\gamma E}$, plus a $G \times S$ vector of daily effort residuals $\boldsymbol{\omega}^{\boldsymbol{E}}$ accounting for day-to-day deviations from the seasonal log means in each stratum. To this sum, a "weekend effect" was added by multiplying an indicator $w_d$ (set to 1 for weekends and holidays) by $\beta_1$, a coefficient which accounted for fixed differences in fishing effort between weekends and holidays, and weekdays (Trudeau et al., 2021). This exponentiated sum was then multiplied by the same indicator variable identifying whether the fishery was open on a particular date $O_{d,s}$:

$$\boldsymbol{\lambda}_{\boldsymbol{d}}^{\boldsymbol{E}} = e^{\mu^{E} + \boldsymbol{\gamma}^{\boldsymbol{E}}\sigma^{\gamma E} + \boldsymbol{\omega}_{\boldsymbol{d}}^{\boldsymbol{E}} + \beta_1 w_d}(o_{d,s}) \tag{5}$$

The daily effort residual $\boldsymbol{\omega}_{\boldsymbol{d}}^{\boldsymbol{E}}$ was modeled in the same manners as the daily catch per unit effort residual (eqs. 2-4).

Total catch $C$ on a given day $d$ for a particular gear type $g$ and river section $s$ was modeled as the product of mean daily catch rate (i.e., CPUE; fish per hour per angler; $\lambda^C$), the mean daily effort (i.e., mean number of anglers, $\lambda^E$), and the total amount of available fishing hours per day (e.g., daylight hours per day, $L_d^D$). Rather than simply multiply these variables, since catch can only be integer values, this product was treated as the rate parameter of a Poisson distribution and daily catch was drawn from this distribution:

$$C_{d,g,s} \sim \text{Poisson}(\lambda_{d,g,s}^C \lambda_{d,g,s}^E L_d^D) \tag{6}$$

***Observation Model***

Although study designs vary, there are generally two main components of the sampling programs; angler interviews and fishing effort counts (see Pollock et al. 1994). Interviews record the number of fish that are caught by anglers for a given amount of angling effort—catch-per-unit-effort (CPUE), which is typically calculated as catch per hour per angler, whereas effort counts either census total fishing effort (anglers × hours) or subsample it in ways that enable extrapolation to total hours fished. Our model was designed to be applied in cases where angler interviews are conducted in conjunction with either direct censuses of effort, or counts of effort surrogates that can, when periodically spatio-temporally paired with effort censuses (to develop bias corrections for the surrogates) facilitate estimation of effort indirectly.

We modeled the catch $c$ of individual anglers or groups $a$ for a given day $d$, gear type $g$ and river section $s$ using Negative Binomial distribution, i.e.:

$$c_{d,g,s,a} \sim \text{Negative Binomial}(\lambda_{d,g,s}^C H_a, r^C) \tag{7}$$

where $\lambda^{C}$is the latent mean hourly catch rate defined above in the process model, $H_a$ (data) is the recorded number the number of hours that individual angler or group reported fishing, and $r^{C}$ is the over-dispersion parameter accounting for among angler (or group) variability in CPUE. We parameterized the model such that the likelihood could either be negative binomial or zero-inflated negative binomial to account for the possibility of a greater than expected number of zero-catch interviews conditional on the estimated mean and overdispersion parameter:

$$c_{d,g,s,a} \sim \begin{cases} \pi_{g,s} + \left(1 - \pi_{g,s}\right) \text{ Negative Binomial}\left(\lambda^{C}_{d,g,s} H_a, r^{C}\right) & \text{if } c = 0 \\ \left(1 - \pi_{g,s}\right) \text{ Negative Binomial}\left(\lambda^{C}_{d,g,s} H_a, r^{C}\right) & \text{if } c > 0 \end{cases} \tag{8}$$

where the probability of excess zeros $\pi$ was specific to a gear type and stratum but did not vary over time.

The second component of the creel data collection process consists of angler counts. Angler counts are intended to characterize fishing pressure by providing a snapshot of angler effort on a given day and time. When an effort count encompasses the entire spatial area of a fishery and anglers are perfectly detected, this survey is referred to as census count. Even if this was a census, the true daily mean angler effort is observed with error because the count is instaneous rather than an average across the day, and because the daily average, which may not be an integer (e.g., 3.2 anglers), may not be precisely observed during an instantaneous count. The number of anglers observed during a census count of effort $E$ at given time *i* within a day *d* for a particular gear type *g* and river section *s* was modeled with Negative Binomial error, parameterized as a Gamma-Poisson mixture:

$$E_{d,g,s,i} \sim \text{ Poisson}(\lambda^{E}_{d,g,s} \varepsilon^{E}_{d,\, g,\, s,i} \gamma^{T}_{g,s}) \tag{9}$$

where $\lambda^E$ was the latent mean hourly effort and $\varepsilon^E$ was the Gamma random variate accounting for overdispersion in the census effort counts due to within-day variability in angler pressure. Effort counts were modeled as a Gamma-Poisson mixture to enable explicitly modeling the effort deviation from the daily mean at the time $i$ of a particular census count which enables comparing the effort at a specific time of day between two methods, such as census effort counts and index effort counts (described later). Finally, a zero-to-one bounded offset, $\gamma_{g,s}^T$, was added to the likelihood. This is an offset supplied as data and is the proportion of effort that was counted in the census effort count by section and gear. Typically, this offset is fixed to 1. However, in cases where complete census counts are not possible, the proportion of area being sampled can be supplied as an offset (e.g., Table 2, Method 2).

The model also accommodated census counts of boats $B$, which is a commonly used alternative to census of anglers:

$$B_{d,g,s,i} \sim \text{Poisson}(\lambda_{d,g,s}^E \varepsilon_{d,\,g,\,s,i}^E p_s^B \gamma_{g,s}^T) \tag{10}$$

which were treated identically to census angler counts except that the effort must be converted from anglers into the expected value for boats by multiplying by the proportion of boats per angler $p_s^B$ estimated from interviews (described later).

As the name implies, index effort counts are designed to characterize angling pressure at a subset of locations within the boundaries of the total fishing area. Index effort counts are often employed because it is typically impractical to census fishing effort during each angler count. Therefore, index effort counts are usually conducted multiple times per day on each survey day while census counts are conducted once per day on a subset of survey days. On the subset of survey days when census counts are conducted, the census count is paired with one of the index counts so that they occur at the same time. To model the relationship between the census effort count and the index effort count on day *d* at time *i*, the same residual $\varepsilon^{E}_{d,\, g,\, s,i}$ from the daily mean effort is applied to both the index and census effort counts.

Surrogate or index effort counts can be conducted in many ways. Anglers can be directly enumerated in the subset of "index" sites (so the count is no longer a census at a point in time) and/or indirectly enumerated via counts of their vehicles, boat trailers, or boats. It may be advantageous to enumerate anglers indirectly if vehicles, boat trailers, or boats are more visible and/or located at a smaller number of possible locations, which would result in a higher proportion of the total angling effort being enumerated. However, if vehicles, boat trailers, or boats are enumerated, these indirect counts need to be converted to an estimate of anglers to facilitate estimating catch. Regardless of whether anglers are directly and/or indirectly enumerated, the relationship between angler effort that occurs in the standardized index reaches relative to angler effort throughout the entire river needs to be quantified. Therefore, index counts of anglers *I*, vehicles *V*, boat trailers *T* at given time *i* within a day *d* for a particular gear type *g* and river section *s* were modeled with Poisson error:

$$I_{d,g,s,i} \sim \text{Poisson}(\lambda^{E}_{d,g,s}\varepsilon^{E}_{d,\, g,\, s,i}p^{E}_{g,s}\gamma^{T}_{g,s}) \tag{11}$$

$$V_{d,g,s,i} \sim \text{Poisson}(\sum_{g=1}^{G}\lambda^{E}_{d,g,s}\varepsilon^{E}_{d,\, g,\, s,i}p^{V}_{g}b^{V}\gamma^{T}_{g=1,s}) \tag{12}$$

$$T_{d,g,s,i} \sim \text{Poisson}(\sum_{g=1}^{G}\lambda^{E}_{d,g,s}\varepsilon^{E}_{d,\, g,\, s,i}p^{T}_{g}b^{T}\gamma^{T}_{g=2,s}) \tag{13}$$

where $p_{g,s}^{E}$ is the estimated proportion of angler effort observed in an index area, which is assumed to be a subset of the total. For vehicle and trailer (index) and boat (census) effort count likelihoods, $p_g^V$ and $p_g^T$ are the estimated proportions of vehicles and trailers per angler, respectively. The proportion of anglers with trailers, vehicles, and boats were assumed to be less than one per angler and consequently $p_g^V$, $p_g^T$ , and $p_g^B$ were estimated using a binomial distribution:

$$V_{a,g} \sim \text{Binomial}(p_g^V, A_{a,g}) \tag{204}$$

$$T_{a,g} \sim \text{Binomial}(p_g^T, A_{a,g}) \tag{215}$$

$$B_{a,g}^{A} \sim \text{Binomial}(p_g^B, A_{a,g}) \tag{226}$$

where $V$, $T$, and $B^A$ are the number of vehicles and trailers that individual angler groups $a$ brought to the river, respectively, and $A$ is the number of anglers that are part of that particular interview group. Typically interview groups for catch (eq. 7-8) are the same interview groups as in eq. 14-16, however indexes are separated to accommodate special cases where either the interview data is incomplete or only a portion is usable, thereby enabling that interview to be usable independently for either eq. 7-8, or 14-16, or both. The bias terms $b^T$ and $b^V$ (eq. 12-13) can be thought of as expansion factors describing how many more or less vehicles and trailers were counted relative to how many anglers were present and the number of trailers and vehicles expected based on the proportions of anglers bringing trailers and vehicles to the river. The bias term can be less than, equal to, or greater than one depending on what proportion of the total number of vehicles and trailers are counted (i.e., observer efficiency) and what proportion of vehicles and trailers belong to anglers versus non-anglers. For example, if 100% of vehicles can be enumerated but angler and non-angler vehicles cannot be differentiated and 30% of the vehicles belong to non-anglers, then the estimated bias term will be less than one. For direct counts of anglers $p_{g,s}^{E}$ can be thought of as the observer efficiency (i.e., the average proportion of anglers that were observed fishing in the index areas relative to the entire river section) and will be between 0 and 1.

### *Prior distributions and MCMC simulations*

Bayesian analyses of state-space models require that prior probability distributions be specified for all parent and global variables (e.g., initial states, variances, fixed effects). In general, prior distributions were designed to be vague (see Table 1), thereby allowing the likelihoods to dominate in determining the posterior. It is worth noting that the choice of certain priors will vary based on the specific attributes of a particular fishery (e.g., mean daily effort and mean daily catch rate).

### *Alternative Study Designs Currently Implemented*

The process models made heavy use of hierarchical structure (random effects and autoregressive timeseries random effects that resulted in shrinkage of CPUE and effort parameters toward their global log means. Such an approach treats date- and stratum- specific parameters as partially dependent (related by temporal adjacency or group membership) rather than independent, thereby enabling poorly identified parameters (e.g., on days with little data) to be informed by better identified parameters (e.g., on days with more data). This means that if data suggest considerable variability among strata and through time, the model is flexible enough to accommodate that variation. However, the model attempts to “shrink” that variation in the absence of evidence to suggest otherwise, consequently reducing the effective parameter count and enabling the model to remain identifiable even in data-poor situations. There are several benefits of this. First, one does not need to observe all states on all dates and in all strata. Second, one may “mix and match” likelihood contributions collected at different frequencies and even irregularly. As of writing, we have used this flexible model structure to analyze data arising from creels with diverse study designs including those where different types and quantities of data are available (Table 2), including studies with and with-out total angler effort census counts and studies with angler index effort counts and those with only boat or car and trailer index effort counts.

### *Transparent Reproducible Science*

All code and data used to fit the described model, including the Skagit River 2021 wild winter steelhead sport fishery example dataset, are in a GitHub repository (https://github.com/tbuehrens/CreelEstimates_publication). The code that loads the data retrieves it directly from www.data.wa.gov, which is the state of Washington’s public dataset repository. Since that is a “live” dataset, archival copies of the exact dataset used in the analysis were saved as .rds objects to enable exact replication of the analysis herein. However, the same data retrieval and model fitting code is available in a separate GitHub repository that is actively developed and maintained by the Fish Program of the Washington Department of Fish and Wildlife to analyze creel survey data on an on-going basis (https://github.com/wdfw-fp/CreelEstimates), facilitating transparency and enabling public replication of analyses used to estimate catch in sport fisheries.

### ***Example of Model Implementation-- 2021 Skagit River steelhead fishery***

The Skagit River, located about two hours north of Seattle, WA, is the largest tributary of Puget Sound, and the location of culturally and economically important fisheries for wild winter steelhead for four Native American tribes as well as sport anglers. Because of prolonged regionwide and local declines in population status, Skagit winter steelhead were listed as Threatened under the Endangered Species Act in 2007 (NOAA 2007), resulting in a multi-year closure of the fishery. In 2017 NOAA Fisheries approved a fishery management plan proposed by Washington State and the four tribes with treaty-reserved fishing rights in the watershed (NOAA 2017).

The federally-approved fishery management plan identified an abundance-based harvest control rule with progressively increasing allowable exploitation rates as wild winter steelhead run sizes increase. The state and tribes annually produce a pre-season forecast of the run size using an Integrated Population Model originally developed by Scheuerell et al. (2021) and adapted to consider alternative recruitment residual and age composition error structures subsequently (Casey Ruff, Swinomish Indian Tribal Community; https://github.com/casruff/Skagit-River-Steelhead-Forecast). This model projected a run size of 4,297 wild winter steelhead in 2021, enabling a sport fishery with an allowable exploitation rate of 10% (Table 3). Tribal fisheries involve harvest of wild steelhead. The sport fishery is mark-selective, meaning that wild steelhead are required to be released, and fishery exploitation occurs only through indirect mortality of released fish. Exploitation in the sport fishery is therefore estimated as catch multiplied by the catch and release mortality rate, which is identified in the fishery management plan to be 10% of wild steelhead caught in the sport fishery (NOAA 2017). The sport fishery is generally allocated approximately 50% of the allowable harvest, with the remainder being allocated to the tribes.

Creel surveys on the Skagit River were initiated on February 1st, 2021, which marked the opening day of the fishery and continued through the end of the season on April 13th, 2021. The fishery was open four days per week (Saturday – Tuesday) on parts of the mainstem Skagit River (from the Dalles Bridge in the town of Concrete up to the Cascade River Road Bridge in Marblemount) and its main tributary, the Sauk River (from the mouth of the Sauk up to the Darrington Bridge) (Figure 1). Sport fishing regulations required catch-and-release for all wild steelhead and no fishing from one hour after sunset to one hour before sunrise.

A stratified random survey design was used to conduct roving-roving creel surveys following the methods outlined in Malvestuto et al. (1978), Pollock et al. (1994) and Hahn et al. (2000). First, the survey was stratified by day-type (weekday or weekend) and four to five sample days were randomly selected consisting of two to three weekdays and both weekend days per week. Second, the survey was stratified within each sample date by shift (AM or PM). The second temporal stratification is typically used to allocate sampling effort to either an AM or PM time frame. However, both “shifts” were sampled on each survey date and the stratification was mainly used for scheduling staff. Specifically, AM shifts began approximately at sunrise and continued for the next six to eight hours on the river while PM shifts began six to eight hours before sunset and lasted until sunset. On each survey date, two technicians were assigned to each shift, and one technician surveyed the mainstem Skagit River and the other surveyed the Sauk River.

During each sample day, technicians conducted both index effort counts and angler interviews. Generally, three index effort counts were conducted per day and the start times were randomly selected. Effort counts were designed to be "instantaneous" in that all index sections of a particular river were surveyed in approximately one hour or less. During each effort count the technician would survey the predefined set of index locations, which were thought to include virtually all areas anglers parked their cars and trailers, and enumerate the number of parked vehicles and boat trailers. In between effort counts, anglers were opportunistically interviewed. When a group of anglers was fishing together, the data for the group were collected as a single interview. During an interview, the angler (group) was asked a predefined list of questions that included: angler type (boat, shore), gear usage (fly, gear), number of anglers in the group, number of vehicles and boat trailers brought to the river, when fishing started and ended, trip status (incomplete, complete), fishing location (Skagit, Sauk), and number of fish caught. If any fish were caught, additional information was collected for each individual fish, including species, origin (hatchery, wild), and fate (harvested, released). Census angler effort counts were conducted by helicopter on a subset of sample days, during which anglers were directly enumerated and classified as either a boat or shore angler.

### ***Resampling data to optimize monitoring programs***

One advantage of developing a fully generative creel model that quantifies uncertainty is the ability to adjust study designs to achieve desired precision. To illustrate an implementation of this, we refit the Skagit 2021 wild winter steelhead fishery dataset to six data subsets that represented realistic tradeoffs managers might navigate to achieve cost efficiency and desired precision while remaining logistically feasible (as opposed to subsampling data randomly in ways that would not be operationally feasible). We simulated through data removal the effect of conducting three levels of census effort count aerial flights: 1) the baseline of 5 flights that covered all river sections, 2) a scenario where only flights 1, 3, and 5 were completed, and a scenario where only one flight (the third) was completed in the middle of the season. We also simulated through removal the effect of conducting interviews and index effort counts daily as occurred, versus conducting interviews and index counts every other day. We simulated scenarios for these two data types simultaneously since the same staff conduct interviews and index effort counts, leading to little cost savings of eliminating one or the other data type independently. We summarized results by quantifying the coefficient of quartile variation for both season total catch and effort, and quantifying bias as a percent difference of an estimate from the estimate made using all of the data (treating this estimate as "truth").

### ***Model Fitting***

The creel state-space model was fit using the software program Stan (Carpenter et al. 2017) to generate samples of the posterior distribution of all unknown parameters in the model. Stan implements gradient-based MCMC methods termed "Hamiltonian Monte-Carlo" via the "No-U-Turn" sampler and often provides improved sampling efficiency that leads to faster inference and less autocorrelation among iterations when compared to other MCMC software implementations (Ames and Au 2018). We interfaced with Stan using Program R (R Development Core Team 2019) and the *rstan* package (Stan Development Team 2018a). We ran four Markov chains with a burn-in period of 1,000 iterations followed by an additional 1,000 iterations where samples were saved for posterior inference. Chains were not thinned because doing so discards information, and because there was little evidence of autocorrelation among draws. Models were accessed for convergence using a range of diagnostic tools offered in the *ShinyStan* package (Stan Development Team 2018b) including, visual examination of posterior distributions and trace plots, calculation of effective sample sizes (ESS) and $\hat{R}$, and checking for divergent transitions (Stan Development Team 2018c). The maximum $\hat{R}$ value was 1.019 and the minimum effective sample size was 247. Based on these criteria and the visualizations, we concluded that our reported posterior distributions are accurate and represent the underlying stationary distributions of the estimated parameters.

Posterior predictive distributions (Gelman et al. 2013) were generated to assess the fit of the model to the data for all observations in all likelihoods that were available in the Skagit dataset. Plots of observations versus 95% posterior predictive intervals (PPI) were generated for each likelihood, separated by strata (boat vs. bank anglers, and river sections) to determine if certain strata were associated with better or worse fit to the data. The proportion of observations that fell outside 95% posterior predictive intervals (PPI) were calculated for each likelihood. Validity of nominal interval coverage is indicated by a match between the PPI coverage and the percentage of data points occurring within the interval. A greater percentage of points occurring within the PPI suggests that the data are consistent with the generative processes in the model, but the model may overestimate uncertainty, whereas a smaller percentage of points occurring within the PPI indicates that the model does not adequately capture variability in the data, particularly if the difference is large (Gelman et al. 2013).

### *Probabilistic Management Metrics*

One utility of using the Bayesian framework including generative models is their inherently robust quantification of uncertainty (e.g., Magnusson et al., 2013). This is useful because uncertainty in a natural resource management context represents risk, and the estimates generated using this modeling approach can therefore be used to quantify risk and manage fisheries in a probabilistic manner to balance opposing risks (e.g., of over-fishing or conversely, unnecessarily foregoing angling opportunity). To illustrate the utility of probabilistic management we propagated the uncertainty in the pre-season forecasted total run size and the creel-based estimate of catch to generate probabilistic estimates of the proportion of fishery impacts used by the 2021 Skagit wild winter steelhead fishery, as well as what proportion of impacts remained available, and illustrated tradeoffs.

The first step in probabilistic management was quantifying the allowable harvest. This was done by multiplying each draw of the run size forecast by the corresponding allowable exploitation rate (Table 3) and expressing the result as a cumulative percent distribution or exceedance probability (i.e., the probability that the allowable harvest exceeded specified number of fish). The second step was to determine what proportion of allowable harvest the sport fishery had used by taking the posterior distribution of total sport catch, multiplying it by the fishery management plan-stipulated 10% assumed catch and release mortality rate (NOAA 2017), and dividing the resultant distribution draw-by-draw by the posterior distribution of allowable harvest. This too could be visualized as an exceedance probability (i.e., the probability that a proportion of the allowable harvest had been used by the sport fishery). Finally, we calculated a) the posterior median estimate of the proportion of total allowable (sport plus treaty tribal) harvest used, as well as the probability that sport fishery harvest had exceed both 50% (approximately the state’s share) and 100% (i.e., the possibility the sporty fishery used both its shared plus the tribal share) of the total allowable harvest under the harvest control rule.

### 3.D. Results & Discussion

A range of innovative approaches has been developed to quantify angler effort and limit the necessity of expensive census counts, largely as an outgrowth of roving creel study designs described several decades ago (e.g., Malvestuto et al. 1978, Pollock et al. 1994). Common approaches include quantifying indices of fishing effort such as counting anglers at a subset of locations, or counting angler surrogates such as boats, vehicles, and boat trailers. Angler interview-reported ratios of anglers to angler-surrogates (e.g., vehicles, trailers, boats) enable estimates of effort generated from surrogates. Such angler index- and surrogate-based effort estimates may be compared with less frequent census effort counts by comparing temporally and spatially synchronized estimates made with the two approaches, in turn enabling quantification of bias in the indirect approaches. Bias estimates can then be applied to indirect estimates of effort when census counts are not available. However, the ability to quantify uncertainty and estimate effort jointly from these disparate data sources and methods has not previously been possible. We demonstrated that such joint estimation and propagation of uncertainty is possible and can be used to successfully estimate effort and catch using our multivariate state space model,

### ***Skagit Steelhead Fishery Data and Goodness of Fit***

The Skagit fishery spanned 72 days in 2021, of which 42 were open for fishing based on a four day per week schedule. Of these, both river sections (Skagit, Sauk) were creel sampled on 41 days, including 3 index effort counts per day and a total of 1,835 angler interviews. Five aerial flight census effort counts occurred and were spread throughout the season in an attempt to measure effort as well generate a representative estimate of the bias-correction parameter that related these counts to the three daily effort index counts.

Posterior predictive checks indicated that the model fit the Skagit River steelhead fishery data well. The model fit was good for both interview and effort count likelihoods, though slightly better for effort counts. Effort count likelihoods, including census angler counts made periodically with helicopter flights, and index effort counts of vehicles and trailers made each day the river was surveyed, all had zero observations that were outside of 95% posterior predictive intervals (Table 2, Figs 2 and 3). Angler group interview likelihoods, including catch, angler group vehicle counts, and angler group trailer counts also fit the data well, but some observations fell outside the 95% PPI; 95% PPI coverage for catch 0.97 (Sauk River boat anglers) to 0.99 (Sauk River bank anglers) depending on the section and angler type. Coverage of 95% PPI for angler group trailer counts ranged from 0.99-1.00, also indicating good fit of the model to the data, however, only 93% of angler group vehicle counts fell within the 95% PPI indicating true interval coverage was slightly below nominal interval coverage (Table 2, Figures 4-5).

Although all likelihood contributions had actual PPI coverage that was very close to, or greater than, nominal, we evaluated alternative model structures to explore model adequacy. In particular, the version of results reported utilized a negative binomial likelihood for angler interview catch, however, we also fit a version with a zero-inflated estimator. PPI coverage for the two estimators was nearly identical; one fewer data point was contained in the 95% PPI for the zero-inflated version. We elected to use the uninflated version due to the similar fit, negligible differences in estimated catch between the models, and a smaller parameter count (more parsimonious). However, some data sets may require the zero-inflated version to ensure adequacy of fit. We did not test a zero-inflated version of likelihoods for census or index effort counts because after examining several datasets including the 2021 Skagit winter steelhead dataset, we determined that the datasets were not zero-inflated (either no, or very few, zeros). Other datasets, however, might require this modification.

### ***Skagit River Steelhead Estimates of Catch, Effort, and CPUE***

The model successfully generated seasonal estimates of total catch, effort, and CPUE. Season total catch in the 2021 Skagit winter steelhead sport fishery was estimated to include 974 (posterior median) wild steelhead caught and released (95% CI = 779-1211) (Table 3, Figure 6). Season total fishing effort in the 2021 sport fishery was estimated to include 35,444 (posterior median) hours of angler fishing effort (95% CI = 30,679-40,684) (Table 3). Season total mean CPUE was 0.028 fish per hour (posterior median) wild steelhead caught and released (95% CI = 0.023-0.032) (Table 3, Figure 6). The CPUE estimate included all angler types and fishery sections and accounted for unequal effort among them, calculated by dividing the posterior distribution of season total catch by the season total of effort. Estimates of total catch and effort were highly precise, indicated by a coefficient of quartile variation of 7.6% for season total catch and 4.8% for season total effort. The coefficient of quartile variation is similar to the coefficient of variation but more robust to skew and outliers (Bonnett 2006). Such high precision was not altogether surprising given high sampling rates in the creel survey; the total number of angler hours reported in interviews was 17,148 and the total of inter-reported catch was 435, which were 48% and 45% of the estimated season totals, respectively.

Catch, effort, and CPUE were considerably lower for bank anglers than boat anglers and also lower for the Sauk River section than the Skagit River section.  Posterior median catch estimates for boat anglers were 512 and 399 fish in the Sauk and Skagit sections, respectively, versus 16 and 41 fish for bank anglers in these river sections (Table 3). Posterior median effort estimates for boat anglers were 12,920 and 12,022 hours in the Sauk and Skagit, respectively, versus 1,823 and 6,918 hours for bank anglers in these river sections. Effort started low for both boat and bank anglers in the Sauk and Skagit but picked up in the third week of the fishery for the Sauk, while slowly increasing over the entire season in the Skagit. CPUE was very low in the Skagit until the middle of March but peaked for boat anglers in late February in the Sauk (Figure 7). Posterior median estimates of seasonal (across days) median CPUE estimated for boat anglers were 0.034 and 0.029 fish per hour in the Sauk and Skagit, respectively, versus 0.006 and 0.008 fish per hour for bank anglers in these river sections (Table 3).

### ***Use of Creel in Probabilistic Management***

The quantification of uncertainty in both (pre-season) forecasts of wild winter steelhead run size and catch in the sport fishery enabled probabilistic management and quantification of opposing risks. The pre-season forecast was a run size of 4,297 wild winter steelhead (posterior median, 95% CI = 2,283-8,166) (Figure 8a). As a result of the forecast probability density function spanning all levels of the exploitation rate matrix, there was a non-zero probability that any of the exploitation rates in the harvest control rule's exploitation rate matrix (Table 3) might apply. Exceedance probabilities calculated using the run size forecast and the exploitation rate matrix suggested that there was nearly a 100% chance of allowable harvest exceeding 100 wild winter steelhead, a greater than 50% chance that allowable harvest would be at least 400 fish, a 50% chance the allowable harvest would be 430 (10% of the median pre-season forecast, corresponding to the second exploitation rate bin; Table 3), and non-negligible probabilities existed that the allowable harvest might be as high as 2,000, with nearly zero probability of greater allowable harvest  (Figure 8b). The approximate sport share of allowable harvest was obtained by multiplying the allowable harvest by 50%. This could then be converted into allowable catch and release encounters by multiplying by 10 since each catch and release event represents 0.1 mortalities based on the management plan (NOAA 2017).

The creel catch estimation model was fitted each week during the season to estimate cumulative catch and calculate the proportion of allowable catch used by the fishery and remaining available. Opposing risks existed because of uncertainty in the run size, which corresponding exploitation rate ceiling would apply in 2021, and in the creel-estimated seasonal cumulative catch, by the planned end of the fishery on April 13, 2021. On one hand, the estimated portion of total allowable harvest used by the sport fishery was 23% (Figure 8c-d), or 46% of the approximately 50% share of total allowable harvest available to sport fisheries (the other 50% is reserved by federal treaty for Native American tribes in the Skagit Basin with treaty-reserved fishing rights), meaning that if managers closed the fishery, they risked foregoing available angling opportunity in a popular fishery. On the other hand, due to the considerable uncertainties described, the probability that sport fishery had already exceeded a 50% share (e.g., risk-neutral management) was 40%, and there was a 4% chance that the sport fishery had exceeded the total allowable harvest for sport and tribal fisheries combined (Figure 8c-d). Managers, however, closed the fishery as scheduled, in effect prioritizing the risk that the fishery had exceeded its share of allowable harvest over the risk that available catch remained unutilized. The model results highlight how future improvements in the precision of forecasts and estimated catch could enable the sport fishery to access a greater portion of available mortalities while not incurring additional risk. Alternatively, future iterations of the management plan could consider implementation error (e.g., Punt et al., 2016), lessening the need implement the harvest control rule without error.

### ***Optimizing study designs for desired precision and accuracy***

Simulated data removal suggested tradeoffs useful for optimizing study design of creel surveys. In general, precision of all target estimates was high with the coefficient of quartile variation (CQV) ranging from 5-13% for season total effort across all scenarios, and from 8-17% for season total catch (Figure 9). This is probably, in part, because nearly 50% of all effort and catch was estimated to have been sampled in the full data scenario, meaning that even in the worst case, 25% of effort and catch would have been creel sampled (using only even-day sampling). Additionally, the modest declines in precision and accuracy by dropping aerial census flights was likely, in part, due to limited variability in the bias parameter relating more frequent index effort estimates with census-based effort estimates. Such insensitivity to a reduction in census effort counts may not occur in other fisheries with more variable bias.

Both precision, measured through the coefficient of quartile variation, and accuracy (relative to estimates using the full dataset) of total catch and effort were more negatively affected by removal of 50% (odd number days) of creel interview and index vehicle and trailer effort count data than removal of all but one out of the five aerial effort counts (Figure 9). This is likely because elimination of creel days both resulted in a loss of interviews and effort counts, whereas loss of flights only affected effort estimates. On the other hand, the costs associated with a small number of flights alone may be comparable to the loss of 50% of interview days due to the high hourly cost of privately charted helicopter flights. Ultimately determining the most desirable study design for the Skagit steelhead fishery would require additional information including managers' desired precision and updated costs for labor and flights. However, the approach presented here facilitates a formal optimization of the study design given those variables and demonstrates the results of varying levels of survey effort.

### *Assumptions*

The accuracy of any estimator is conditional upon satisfying the assumptions implicit in the estimator. Since unbiased estimates is the chief objective of most catch monitoring programs, identification and examination of underlying assumptions is an important part of monitoring programs. In addition to the data fitting the model (previously reviewed), the structure of the model we developed relied on several additional assumptions: CPUE—1) anglers interviews were an *iid* representative sample with respect to their CPUE within a gear (boat/bank), section, and date stratum, and 2) anglers accurately identified and reported their catch and hours fishing; Effort: 1) surveyors could accurately (and completely) census all anglers during helicopter flights, 2) vehicle and trailer counts were an unbiased measure of effort when adjusted by interview-reported ratios of anglers to vehicles and trailers and a gear type (boat/bank) bias correction factor, 3) the bias correction factor and ratios of vehicles and trailers per angler varied by gear type but did not vary across sections or over time, 4) the timing of index effort counts was randomized but the mean of the daily counts gave an unbiased measure of the daily mean effort, and 5) index effort counts that occurred at the time of the census effort count flights were synchronized and therefore measuring the same level of effort.

Of these assumptions, we believe that most are likely to be met a vast majority of the time. However, two assumptions are worth particular attention: 1) effort should be made to ensure creel samplers move throughout the sections so that all anglers have an equal chance of being interviewed to satisfy the *iid* assumption within a stratum; 2) the invariant bias correction factor for relating index vehicle and trailer counts to census effort counts may be inappropriate if vehicle and trailer counts incorporate variable proportions of all angler vehicles and trailers among sections, or if there is seasonal variability (e.g., because counts include seasonally variable proportions of non-angler vehicles and trailers). In cases where it is possible to census all or nearly all angler cars and trailers and no or only a small number of non-angler vehicles or trailers are present, the assumptions required for the constant bias-correction factor are met. However, the parameterization would need to be adjusted (e.g., to incorporate spatial and temporal variability) if these conditions are not met. A requirement for such spatio-temporal variability in this parameter would be costly in terms of precision and would likely require shrinkage estimators since bias correction parameters are only independently identifiable on days (and at times) when both census and index effort counts occur.

### ***Conclusions***

We developed a flexible multivariate state space creel model that quantifies fishing effort and catch using common roving-roving creel data. We demonstrated that the model adequately fit the data in an example fishery, illustrated the utility of its estimates of uncertainty for managing fisheries to balance opposing risks, and for optimizing field study designs to achieve objectives while managing costs. Widespread application of the methods and software described here will benefit fisheries management and conservation by enabling unbiased estimates of catch and robust quantification of uncertainty across diverse sampling programs.

## 3.E. Acknowledgements

This work would not have been possible without countless hours of angler interviews, and car, boat, trailer, and angler counts by WDFW staff. Funding was mostly provided by Washington State license buyers and taxpayers. Special thanks go to the Columbia River Salmon and Steelhead Endorsement board for funding the initial creels that inspired the development of new methods. Thanks to Danny Warren for building the data systems behind the scenes that make this project work.

### 3.F. References

Ames, A.J., and C. H. Au. 2018. Using Stan for Item Response Theory Models. Meas. Interdiscip. Res. Perspect. 16(2): 129–134.

Bonett, D. G. 2006. Confidence interval for a coefficient of quartile variation. Computational Statistics & Data Analysis. 50 (11): 2953–2957.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell. 2017. Stan: A Probabilistic Programming Language. J. Stat. Softw. 76(1): 1-32.

Clark, J.S., and O. N. Bjørnstad. 2004. Population Time Series: Process Variability, Observation Errors, Missing Values, Lags, and Hidden States. Ecology 85(11): 3140–3150.

De Valpine, P. 2002. Review of methods for fitting time-series models with process and observation error and likelihood calculations for nonlinear, non-gaussian state-space models. Bull. Mar. Sci. 70(2): 455–471.

De Valpine, P. 2003. Better Inferences from Population-Dynamics Experiments Using Monte Carlo State-Space Likelihood Methods. Ecology 84(11): 3064–3077.

Gelman, A., J. B. Carlin., H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2013. Bayesian Data Analysis, Third Edition. New York: Chapman and Hall/CRC. DOI: 10.1201/b1601811.

Hahn, P., S. Zeylmaker, and S. Boner. 2000. WDFW methods manual - Creel information from sport fisheries. Wash. Dep. Fish Wildl. Fish Program Div. Olymp. Wash. Technical report 93-18.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall.

Magnusson, A., A. E. Punt, and R. Hilborn. 2013. Measuring uncertainty in fisheries stock assessment: the delta method, bootstrap, and MCMC. Fish and Fisheries, 14(3): 325-342.

Malvestuto, S.P., W. D. Davies, and W. L. Shelton. 1978. An Evaluation of the Roving Creel Survey with Nonuniform Probability Sampling. Trans. Am. Fish. Soc. 107(2): 255–262.

McCormick, J. L. and J. W. Heckel IV. 2023. Using state-space models to estimate recreational angling effort and infer processes that regulate angler dynamics. Transactions of the American Fisheries Society, 152 (6): 738-748.

NOAA. 2007. Endangered and Threatened Species: Final Listing Determination for Puget Sound Steelhead. 72 FR, pp. 26722-26735.

NOAA. 2017. 4(d) Rule Limit 6 Final Evaluation and Recommended Determination: Skagit River Steelhead Fishery Resource Management Plan. WCR-2017-7053.

Patterson, T.A., L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos. 2008. State–space models of individual animal movement. Trends Ecol. Evol. 23(2): 87–94. doi:10.1016/j.tree.2007.10.009.

Pollock, K.H., C. M. Jones, and T. L. Brown. 1994. Angler survey methods and their applications in fisheries management. Am. Fish. Soc. Spec. Publ. USA.

Punt, A. E., D. S. Butterworth, C. L. de Moor, J. A. A. De Oliveira, and M. Haddon. 2016. Management strategy evaluation: best practices. Fish and Fisheries, 17, 303-334.

R Development Core Team. 2019. R: A language and environment for statistical computing. R Found. Stat. Comput. Vienna Austria. Available from http://www.r-project.org.

Scheuerell, M.D., C. P. Ruff, J. H. Anderson, and E. M. Beamer. 2021. An integrated population model for estimating the relative effects of natural and anthropogenic factors on a threatened population of steelhead trout. Canadian Journal of Fisheries and Aquatic Sciences, 58 (1): 114-124.

Stan Development Team. 2018a. RStan: the R interface to Stan. R package version 2.17.3. Available from http://mc-stan.org/.

Stan Development Team. 2018b. ShinyStan: Interactive Visual and Numerical Diagnostics and Posterior Analysis for Bayesian Models. R package version 2.5.0. Available from http://mc-stan.org.

Stan Development Team. 2018c. Stan Modeling Language Users Guide and Reference Manual, Version 2.18.0. Available from http://mc-stan.org/.

Stobutzki, I. C., M. Miller, and D. Brewer. 2001. Sustainability of fishery bycatch: a process for assessing highly diverse and numerous bycatch. Environ Conserv 28: 167–181.

Trudeau, A., C. J. Dassow, C. T. Solomon, C. M. Iwicki, B. T. van Poorten, S. E. Jones, O. P. Jensen. 2021. Estimating fishing effort across the landscape: A spatially extensive approach using models to integrate multiple data sources. Fisheries Research 233 (2021) 105768.

Tucker, C. M., S. Collier, G. Legault, G.E. Morgan, and D. K. de Kerckove. 2024. Estimating angler effort and catch from a winter recreational fishery using a novel Bayesian methodology to integrate multiple sources of creel survey data. Fisheries Research, 272 (106932).

## 3.G. Tables

Table 1. Prior distributions for model parameters

| Parameter | Default Prior | Notes |
|---|---|---|
| $\boldsymbol{\omega}^{\boldsymbol{C}}_{\boldsymbol{d=0}}$ | $N\left(0, \sqrt{\frac{{\sigma^{C}}^{2}}{1-{\varphi^{C}}^{2}}}\right)$ | sd of this prior is the sd of an AR1 process |
| $\boldsymbol{\omega}^{\boldsymbol{E}}_{\boldsymbol{d=0}}$ | $N\left(0, \sqrt{\frac{{\sigma^{E}}^{2}}{1-{\varphi^{E}}^{2}}}\right)$ | sd of this prior is the sd of an AR1 process |
| $\mu^{C}_{g,s}$ | $N(\log(0.05), 1.5)$ | Vague prior accommodating most salmon and steelhead fishery catch rates; adjust as appropriate |
| $\mu^{E}_{g,s}$ | $N(\log(10), 2)$ | Vague prior accommodating hourly efforts typically seen on river sections of small fisheries; adjust as appropriate |
| $\varphi^{C}$ | $U(-1, 1)$ | Parameterized as: $\varphi^{C} = 2\varphi^{C\ scaled} - 1$, where $2\varphi^{C\ scaled} \sim beta(1,1)$ |
| $\varphi^{E}$ | $U(-1, 1)$ | Parameterized as: $\varphi^{E} = 2\varphi^{E\ scaled} - 1$, where $2\varphi^{E\ scaled} \sim beta(1,1)$ |
| $\sigma^{C}$ | $half - Cauchy(0, 1)$ | Regularizing |
| $\sigma^{E}$ | $half - Cauchy(0, 1)$ | Regularizing |
| $r^{C}$ | ${r^{C}}^{-2} \sim half - Cauchy(0, 1)$ | Regularizing |
| $r^{E}$ | ${r^{E}}^{-2} \sim half - Cauchy(0, 1)$ | Regularizing |
| $\pi_{g,s}$ | $beta(1,1)$ | Vague |
| $b^{\mathrm{V}}$ | $lognormal(0, 1)$ | Centered on zero bias, vague |
| $b^{\mathrm{T}}$ | $lognormal(0, 1)$ | Centered on zero bias, vague |
| $\beta_{1}$ | $N(0,5)$ | Vague |
| $p^{E}_{g,s}$ | $beta(0.5, 0.5)$ | Vague |
| $p^{V}_{g}$ | $beta(0.5, 0.5)$ | Vague |
| $p^{T}_{g}$ | $beta(0.5, 0.5)$ | Vague |
| $p^{B}_{g}$ | $beta(0.5, 0.5)$ | Vague |
| $\boldsymbol{L}^{\boldsymbol{C}}$ | lkj(1) | Hierarchical prior on correlation matrices |
| $\boldsymbol{L}^{\boldsymbol{E}}$ | lkj(1) | Hierarchical prior on correlation matrices |

Table 2. Study design alternatives that have already been successfully implemented using the model presented here. Examples are presented for the sake of illustration; for brevity only the Skagit steelhead fishery analysis is presented herein.

| Data Type | Method 1: (Default; Skagit steelhead) | Method 2: (No effort census counts) | Method 3 (Drano Lake) | Method 4 (Lower Columbia River steelhead) |
|---|---|---|---|---|
| angler group catch | X | X | X | X |
| angler group hrs* | X | X | X | X |
| angler group cars | X | X |  |  |
| angler group trailers | X | X |  |  |
| angler group boats |  |  | X |  |
| angler group (angler count) | X | X | X | X |
| census effort (boat anglers) | X |  |  | X |
| census effort (bank anglers) | X |  | X | X |
| census effort (boats) |  |  | X |  |
| index effort (boat anglers) |  |  |  | X |
| index effort (bank anglers) |  |  |  | X |
| index effort (cars) | X | X |  |  |
| index effort (trailers) | X | X |  |  |

*used in angler group catch likelihood as offset

Table 3. Harvest control rule used by Washington State and Tribal Co-managers in 2021 Skagit wild winter steelhead fishery.

| Run size | Maximum Exploitation Rate |
|---|---:|
| 0-4000 | 0.04 |
| 4001-6000 | 0.10 |
| 6001-8000 | 0.20 |
| >8000 | 0.25 |

Table 4. The proportion of observations that fell within the 95% Posterior Predictive Interval (PPI) for each likelihood contribution, separated by gear and section strata and combined.

| Parameter | Section | Gear | Included | Non included | Total | 95% PPI coverage |
|---|---|---|---|---|---|---|
| Vehicle Counts | Sauk River | NA | 122 | 0 | 122 | 1.00 |
| Vehicle Counts | Skagit River | NA | 121 | 0 | 121 | 1.00 |
| Vehicle Counts | Total | NA | 243 | 0 | 243 | 1.00 |
| Trailer Counts | Sauk River | NA | 122 | 0 | 122 | 1.00 |
| Trailer Counts | Skagit River | NA | 121 | 0 | 121 | 1.00 |
| Trailer Counts | Total | NA | 243 | 0 | 243 | 1.00 |
| Angler Counts | Sauk River | Bank Anglers | 5 | 0 | 5 | 1.00 |
| Angler Counts | Sauk River | Boat Anglers | 5 | 0 | 5 | 1.00 |
| Angler Counts | Skagit River | Bank Anglers | 5 | 0 | 5 | 1.00 |
| Angler Counts | Skagit River | Boat Anglers | 5 | 0 | 5 | 1.00 |
| Angler Counts | Total | Total | 20 | 0 | 20 | 1.00 |
| Catch | Sauk River | Bank Anglers | 680 | 8 | 688 | 0.99 |
| Catch | Sauk River | Boat Anglers | 458 | 12 | 470 | 0.97 |
| Catch | Skagit River | Bank Anglers | 257 | 6 | 263 | 0.98 |
| Catch | Skagit River | Boat Anglers | 404 | 10 | 414 | 0.98 |
| Catch | Total | Total | 1799 | 36 | 1835 | 0.98 |
| Vehicles (Angler Group) | NA | Bank Anglers | 880 | 71 | 951 | 0.93 |
| Vehicles (Angler Group) | NA | Boat Anglers | 820 | 64 | 884 | 0.93 |
| Vehicles (Angler Group) | NA | Total | 1700 | 135 | 1835 | 0.93 |
| Trailers (Angler Group) | NA | Bank Anglers | 939 | 12 | 951 | 0.99 |
| Trailers (Angler Group) | NA | Boat Anglers | 883 | 1 | 884 | 1.00 |
| Trailers (Angler Group) | NA | Total | 1822 | 13 | 1835 | 0.99 |

Table 5. Summary statistics for season-total catch, effort, and CPUE estimates, both aggregated across strata, and split out by section and gear.

| Parameter | section | gear | mean | sd | 2.50% | 25% | 50% | 75% | 97.50% |
|---|---|---|---|---|---|---|---|---|---|
| Catch | Total | Total | 981 | 110 | 779 | 904 | 974 | 1052 | 1211 |
| Catch | Skagit | Bank | 17 | 7 | 6 | 12 | 16 | 21 | 35 |
| Catch | Sauk | Bank | 42 | 13 | 21 | 33 | 41 | 51 | 72 |
| Catch | Skagit | Boat | 404 | 65 | 289 | 358 | 399 | 443 | 544 |
| Catch | Sauk | Boat | 518 | 74 | 387 | 465 | 512 | 566 | 671 |
| Effort | Total | Total | 35505 | 2524 | 30679 | 33751 | 35444 | 37163 | 40684 |
| Effort | Skagit | Bank | 1852 | 373 | 1203 | 1595 | 1823 | 2082 | 2666 |
| Effort | Sauk | Bank | 6982 | 989 | 5188 | 6291 | 6918 | 7639 | 9058 |
| Effort | Skagit | Boat | 12829 | 1172 | 10656 | 12022 | 12778 | 13584 | 15218 |
| Effort | Sauk | Boat | 13841 | 1326 | 11379 | 12920 | 13807 | 14692 | 16613 |
| CPUE (season mean) | Total | Total | 0.028 | 0.002 | 0.023 | 0.026 | 0.027 | 0.029 | 0.032 |
| CPUE (season median) | Skagit | Bank | 0.008 | 0.003 | 0.003 | 0.006 | 0.008 | 0.010 | 0.016 |
| CPUE (season median) | Skagit | Boat | 0.029 | 0.008 | 0.014 | 0.024 | 0.029 | 0.033 | 0.047 |
| CPUE (season median) | Sauk | Bank | 0.006 | 0.003 | 0.003 | 0.005 | 0.006 | 0.007 | 0.013 |
| CPUE (season median) | Sauk | Boat | 0.035 | 0.011 | 0.017 | 0.030 | 0.034 | 0.038 | 0.057 |

## 3.H. Figures

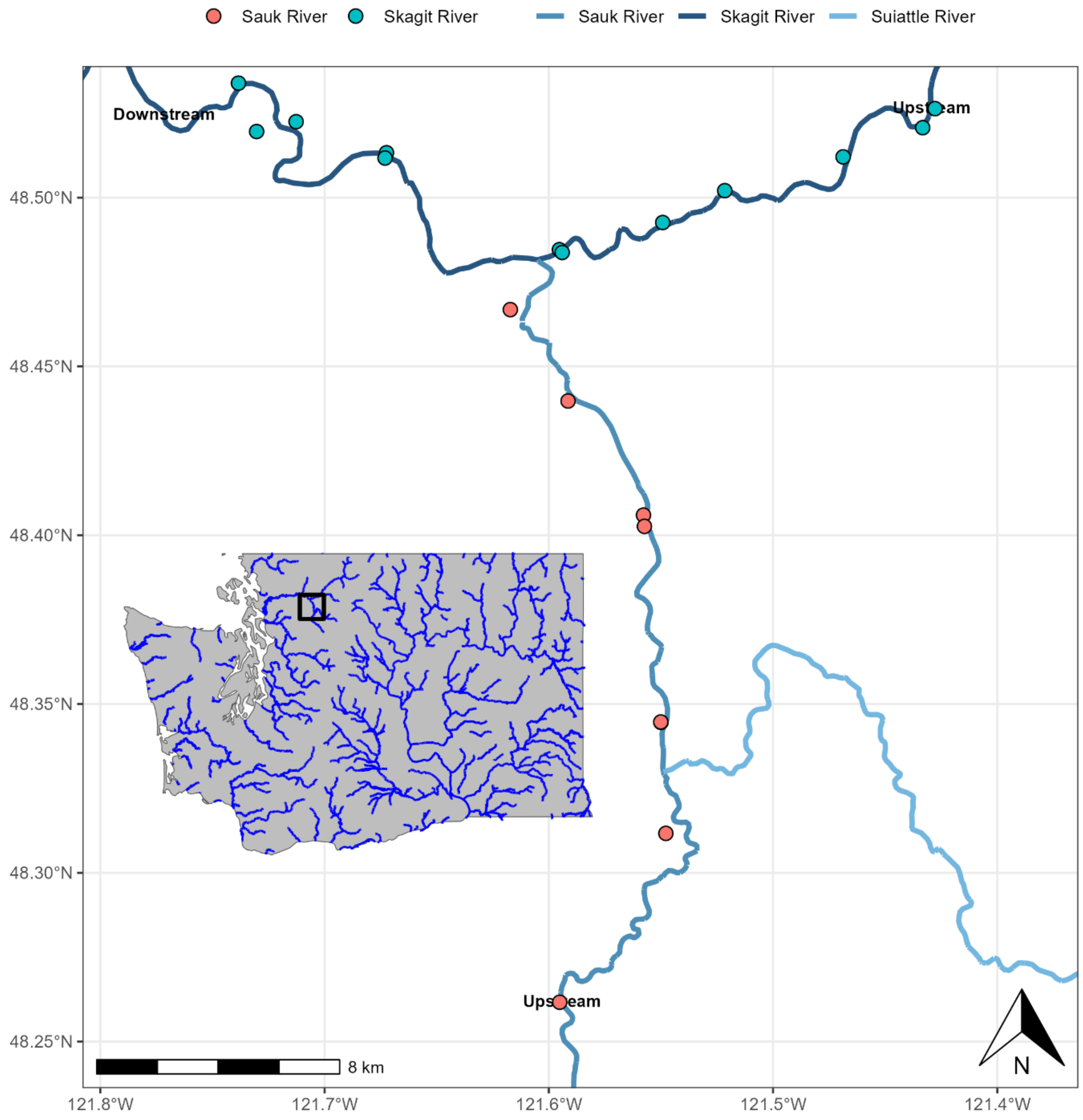


Figure 1. Site map for 2021 Skagit Steelhead fishery showing the locations of the spatial strata (sections; Sauk, Skagit), and index effort (vehicle and trailer) count sites (dots). Upper and lower fishing boundaries on both rivers are denoted by “downstream” and “upstream”. The Suiattle River, a right-bank tributary of the Sauk River, was not open to fishing.

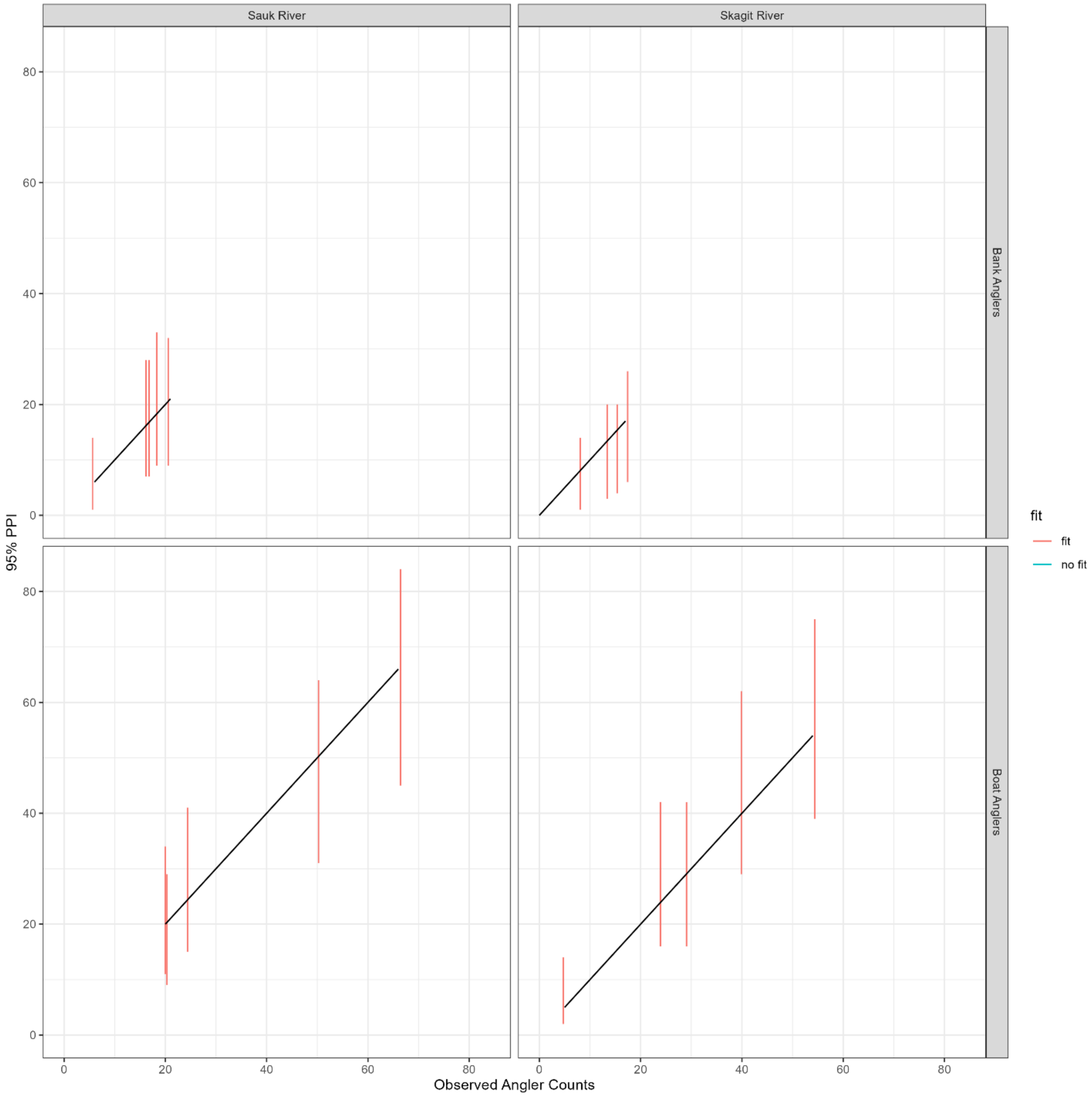


Figure 2. Plots of observed angler counts vs. their corresponding 95% posterior predictive intervals (PPI). A 1:1 line is shown to identify perfect matches between observed and expected. Coloring or PPI indicates which intervals overlap the 1:1 line (red) and do not (blue). A jitter was added to observation values to reduce overlap of identical data points.

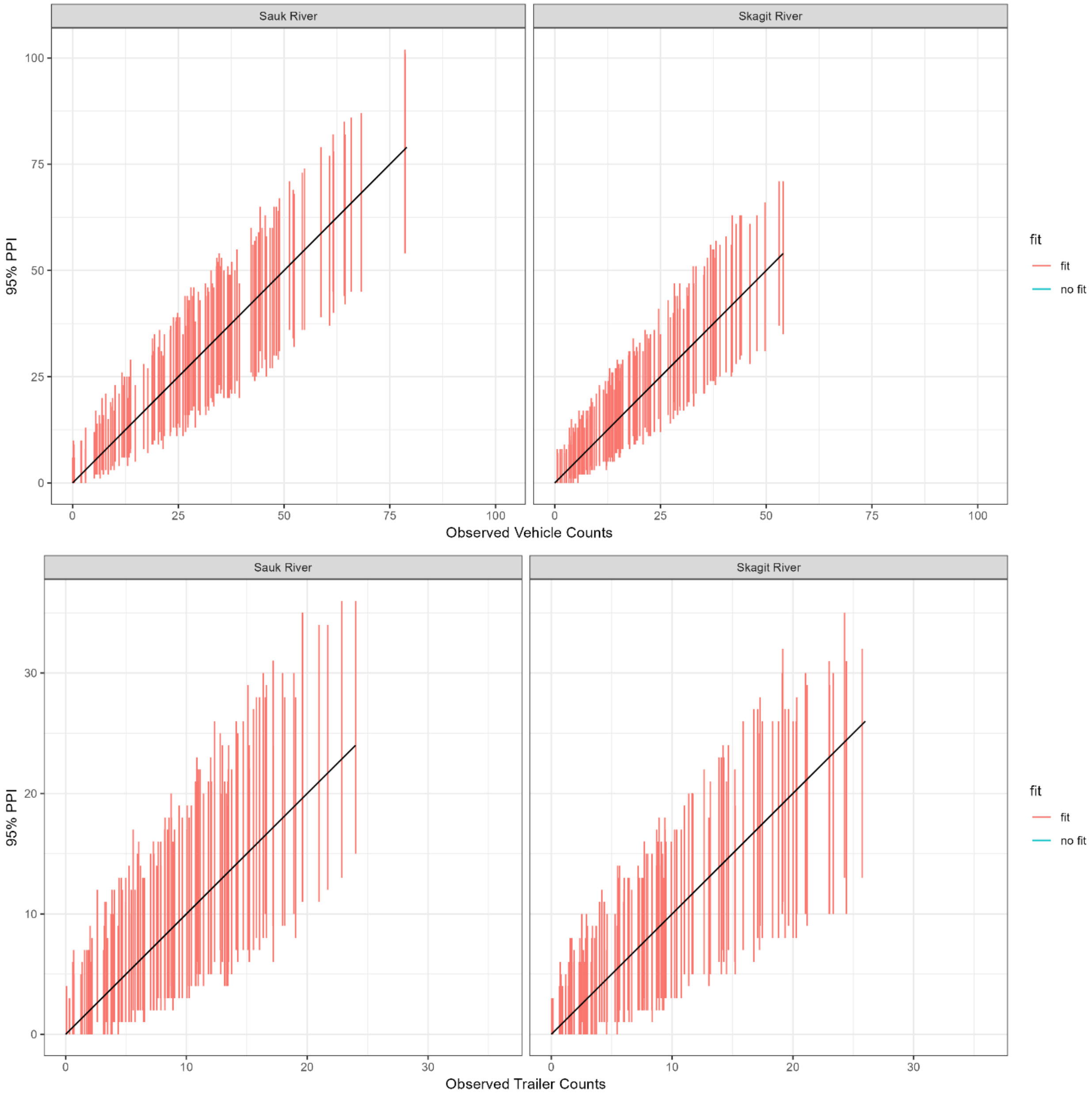


Figure 3. Plots of observed vehicle and trailer counts vs. their corresponding 95% posterior predictive intervals (PPI). A 1:1 line is shown to identify perfect matches between observed and expected. Coloring or PPI indicates which intervals overlap the 1:1 line (red) and do not (blue). A jitter was added to observation values to reduce overlap of identical data points.

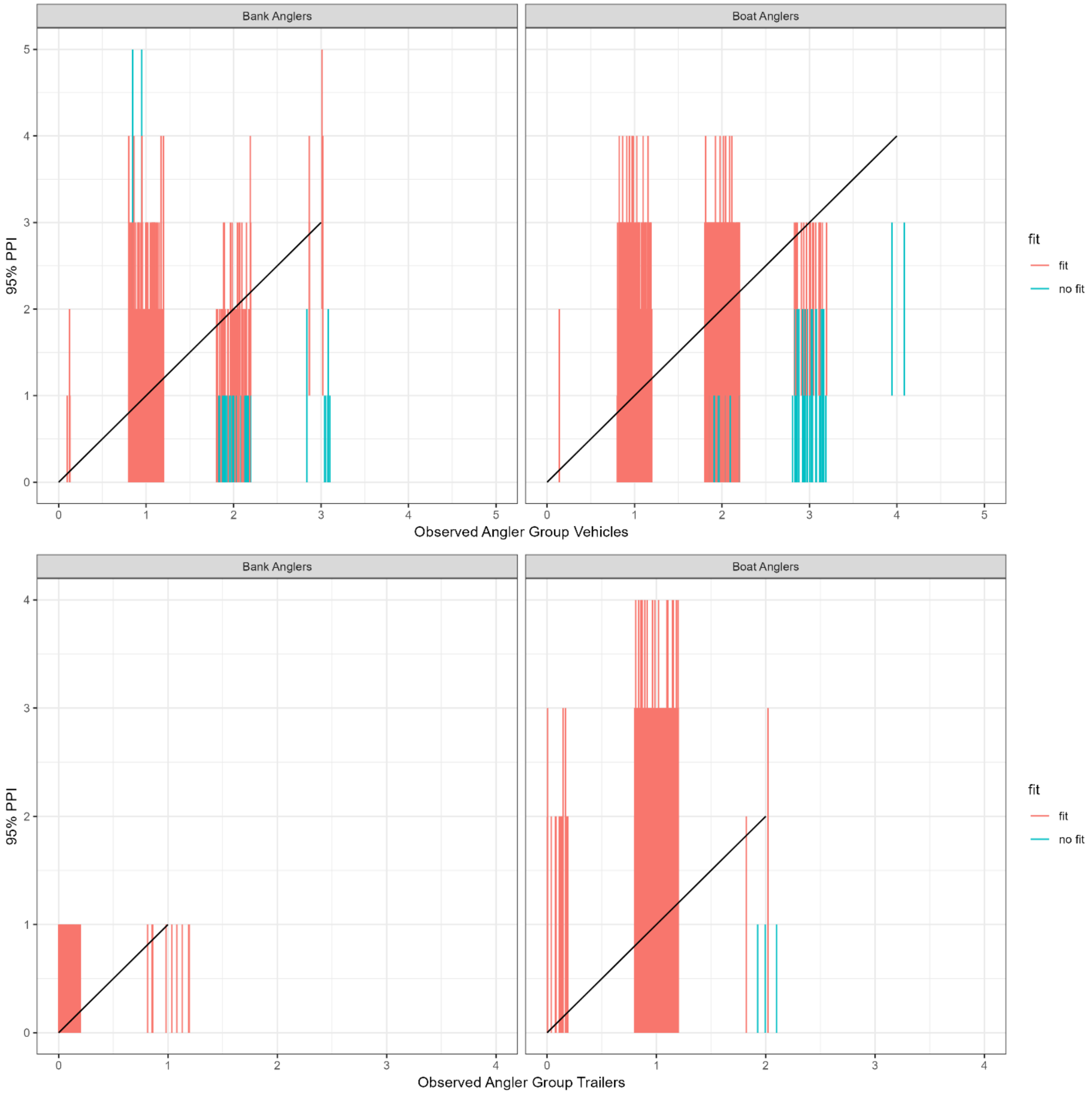


Figure 4. Plots of observed angler group vehicle and trailer counts vs. their corresponding 95% posterior predictive intervals (PPI). A 1:1 line is shown to identify perfect matches between observed and expected. Coloring or PPI indicates which intervals overlap the 1:1 line (red) and do not (blue). A jitter was added to observation values to reduce overlap of identical data points.

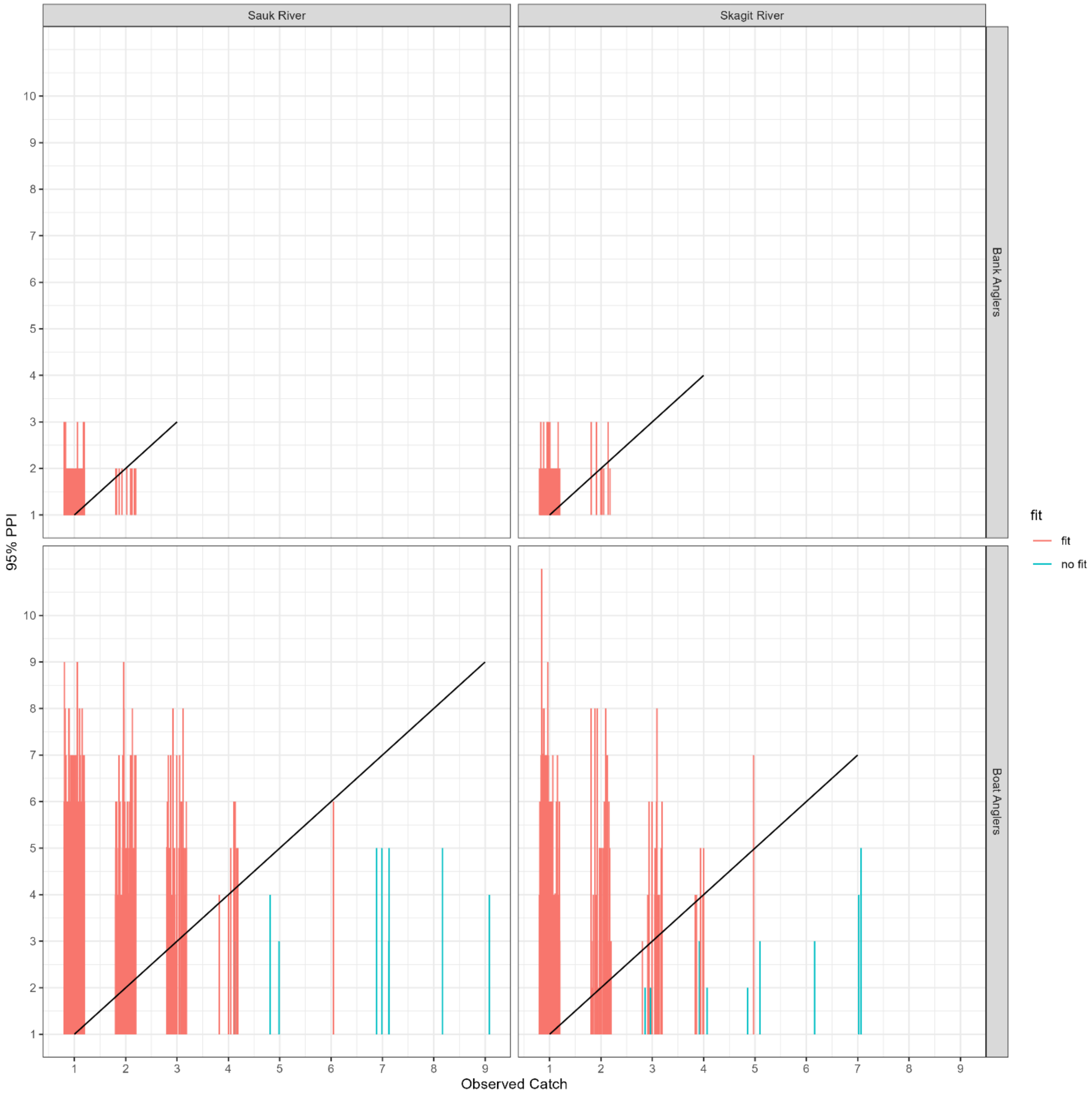


Figure 5. Plots of observed catches vs. their corresponding 95% posterior predictive intervals (PPI). A 1:1 line is shown to identify perfect matches between observed and expected. Coloring or PPI indicates which intervals overlap the 1:1 line (red) and do not (blue). A jitter was added to observation values to reduce overlap of identical data points.

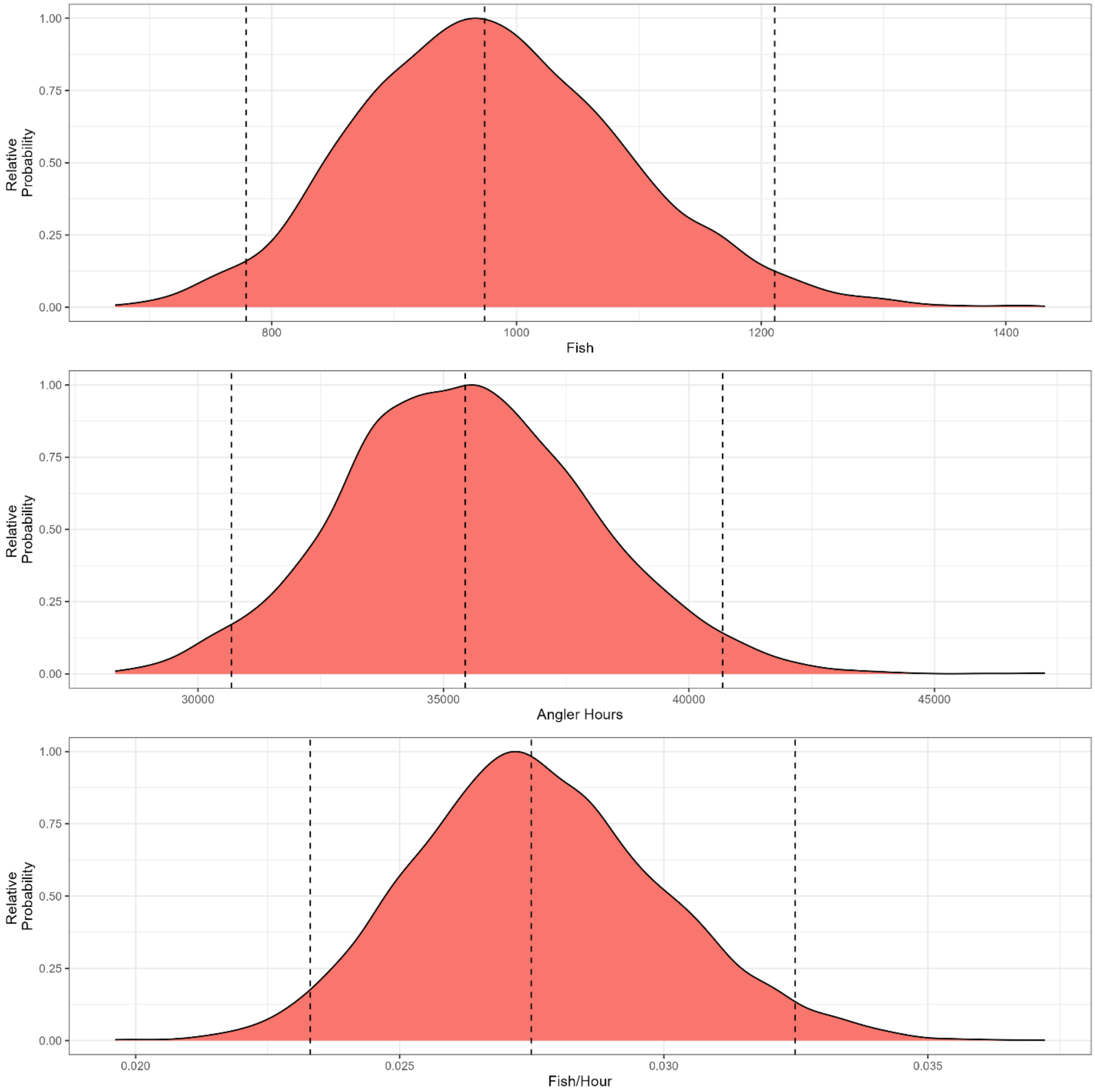


Figure 6. Catch (top) and effort (middle) and seasonal mean CPUE (bottom) probability density plots. Probabilities are standardized relative to the estimates with the greatest probability.

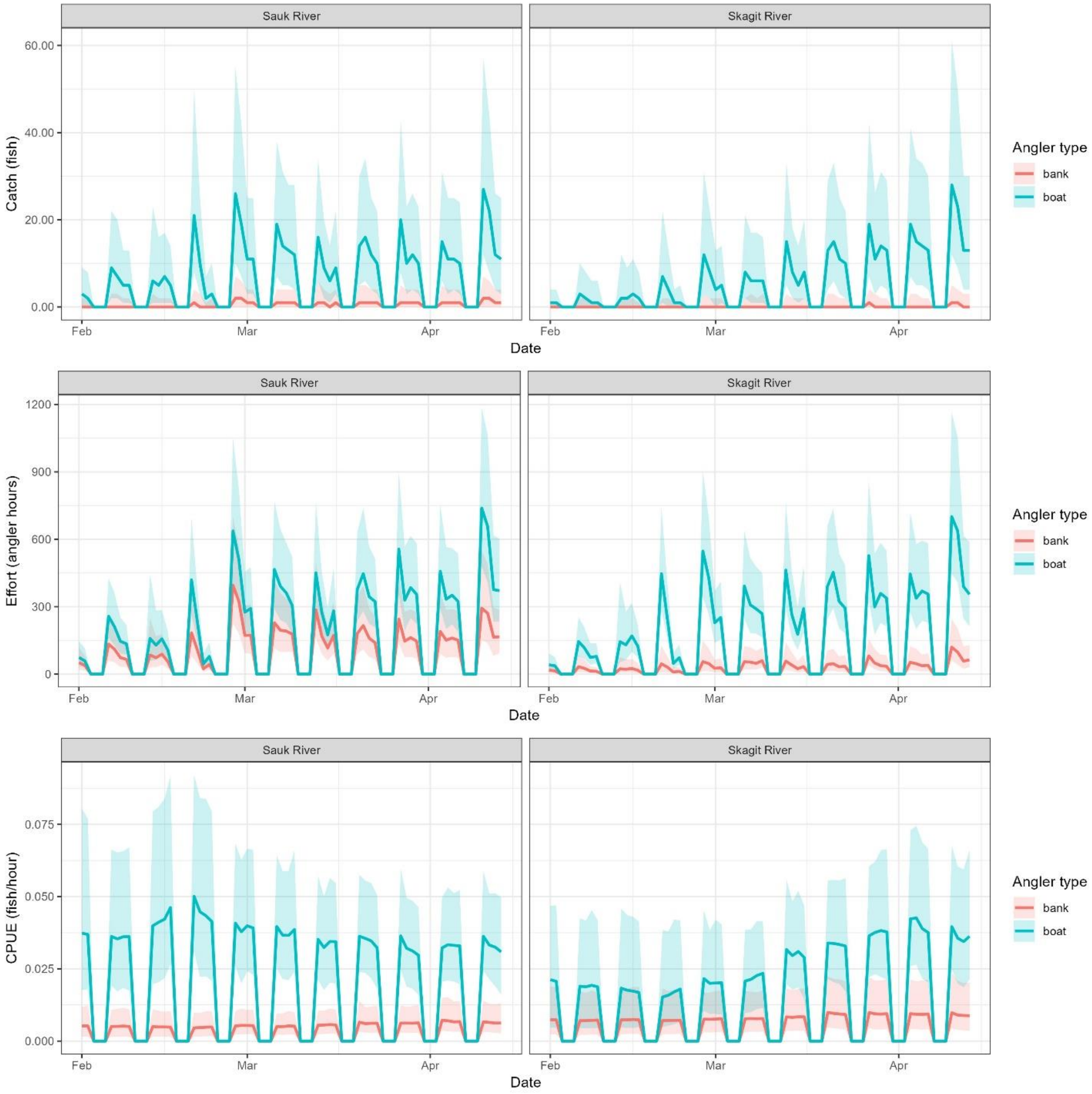


Figure 7. Daily estimates of catch (top), effort (middle) and CPUE (bottom). Lines are posterior medians and shading shows 95% credible intervals. The cyclical weekly pattern is a result of the fishery opening only four day per week.

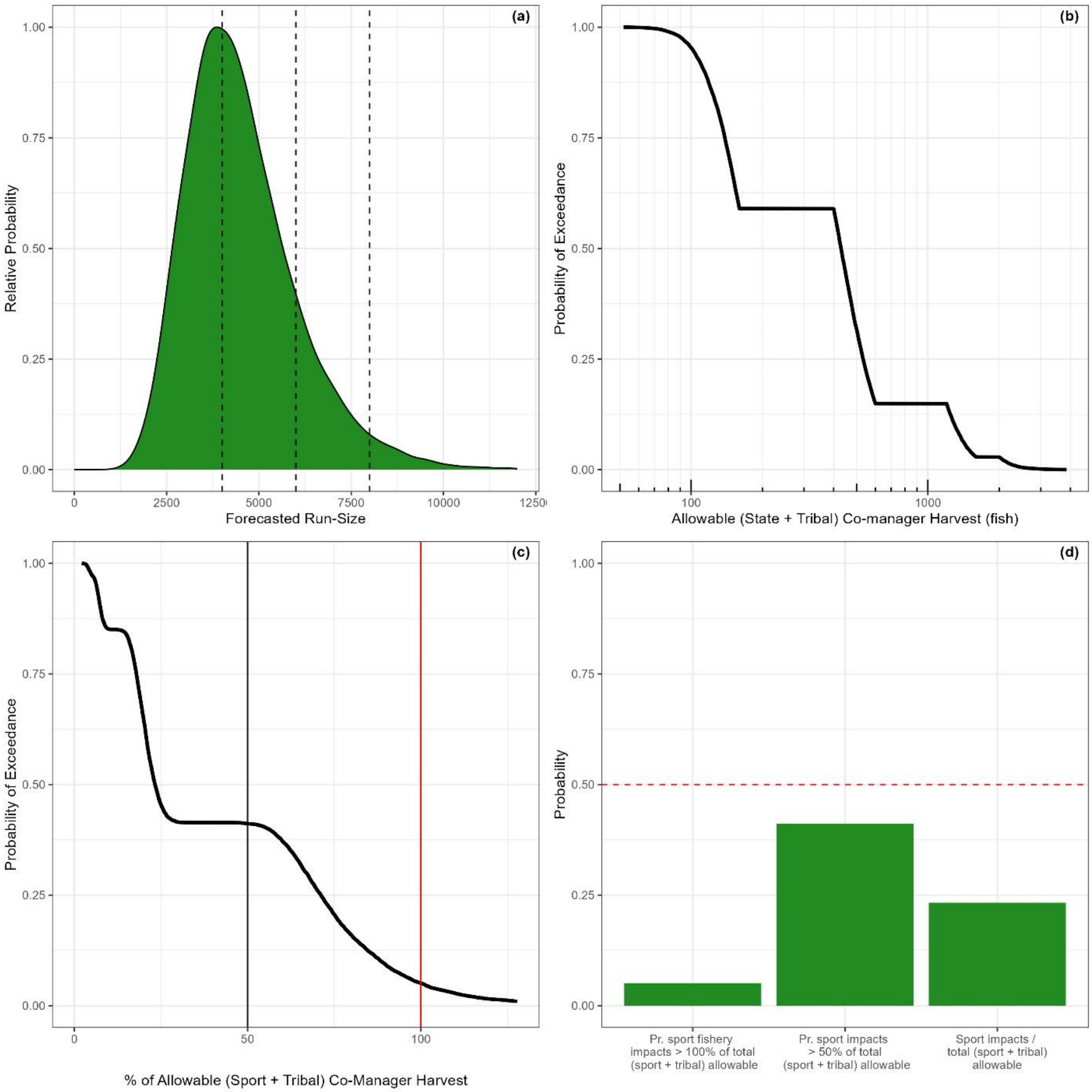


Figure 8. Using creel estimates for probabilistic management of the 2021 Skagit winter steelhead fishery: a) Posterior probability density plot for the forecasted run-size based on methods from Scheuerell et al. 2021, available here: *https://github.com/casruff/Skagit-River-Steelhead-Forecast*. Vertical dotted lines correspond to exploitation rate break points in the harvest control rule (Table 3; NOAA 2017); b) estimated allowable catch exceedance probabilities by applying the harvest control rule to the runsize forecast. Note log10 x-axis; c) the probability that the season-total sport fishery catch exceeded the allowable harvest. Vertical lines denote 50% (an approximation of the non-tribal share) and 100% of the allowable harvest, d) The probability that

the sport fishery exceeded the total allowable harvest, the probability that sport fishery impacts exceeded 50% of the total, and the median estimate of the portion of allowable sport fishery impacts used at the close of the season.

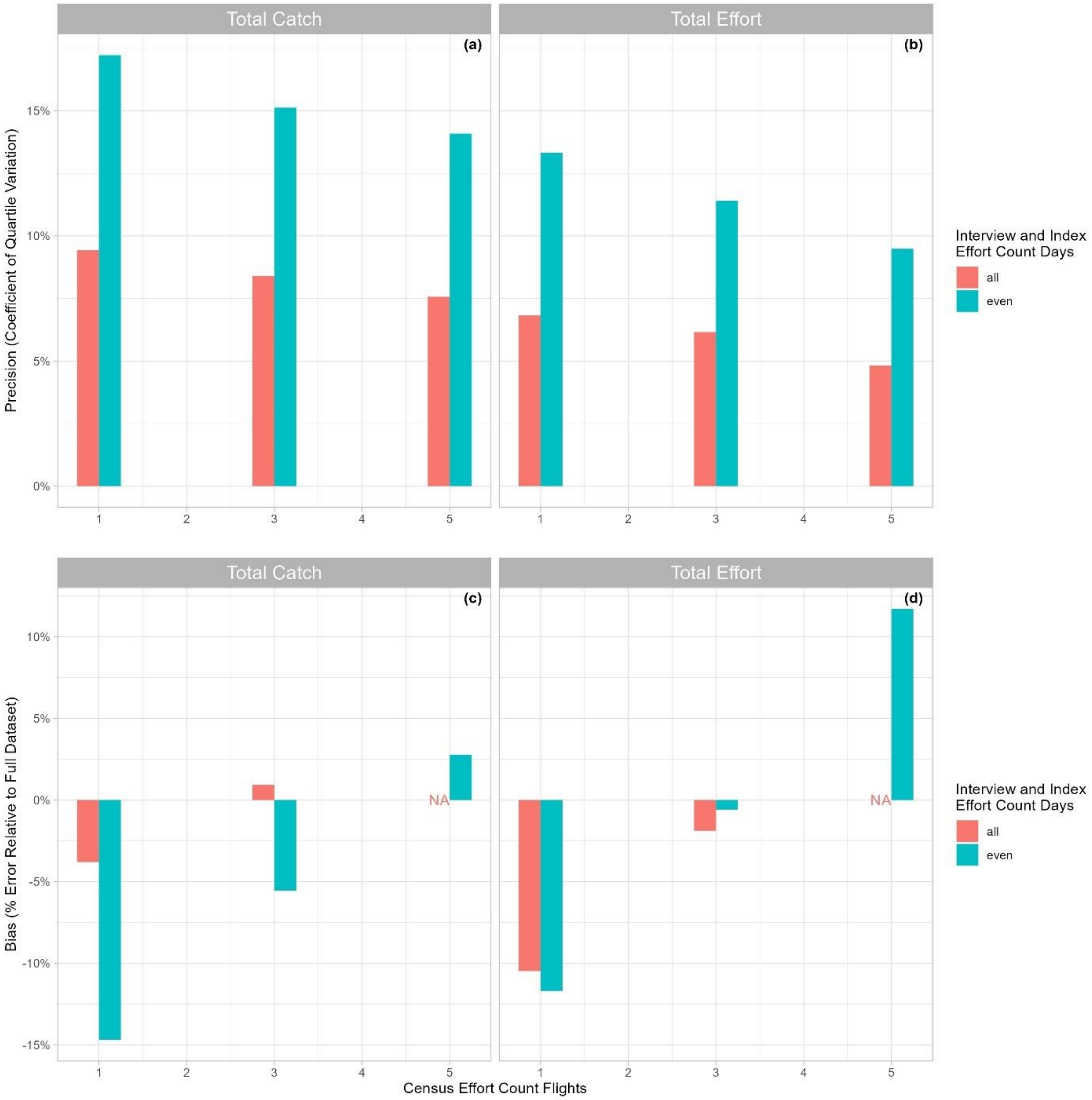


Figure 9. Sensitivity of precision and accuracy of season total catch and effort to data quantity. The model was fit with three levels of census effort count aerial flights: 1) all five flights, 2) only the first, third, and fifth flight, and 3) using only the third (middle) flight. These combinations were pared with two levels of creel effort (interviews and daily index car and trailer effort counts); using all creel data vs. only using data from the even-numbered days. Precision (panels a and b) was calculated as the coefficient of quartile variation (Bonnet 2006), while bias was calculated as the percent difference between the posterior median of an estimate made with a data subset and the posterior median estimate with all data.