

1 Differential impacts of freshwater and marine covariates on wild and hatchery Chinook salmon
2 marine survival

3

4 Brandon Chasco¹, Brian Burke², Lisa Crozier², Rich Zabel²

5 ¹ Fish Ecology Division, National Marine Fisheries Service, National Oceanic and Atmospheric
6 Association, 2032 Marine Dr. SE, Newport, 97365, USA

7 ² Fish Ecology Division, National Marine Fisheries Service, National Oceanic and Atmospheric
8 Association, 2725 Montlake Blvd E, Seattle, USA

9

10 Short title: Integrated marine and freshwater effects on salmon survival

Abstract

Large-scale atmospheric conditions in the Northeast Pacific Ocean affect both the freshwater environment in the Columbia River Basin and marine conditions along the coasts of Oregon, Washington, and British Columbia, resulting in correlated conditions in the two environments. For migrating species, such as salmonids that move through multiple habitats, these correlations can amplify the impact of good or poor physical conditions on growth and survival, as movements among habitats may not alleviate effects of anomalous conditions. Unfortunately, identifying the mechanistic drivers of salmon survival in space and time is hindered by these cross-habitat correlations. To address this issue, we modeled the marine survival of Snake River spring/summer Chinook salmon with multiple indices of the marine environment and an explicit treatment of the effect of arrival timing from freshwater to the ocean, and found that both habitats contribute to marine survival rates. We show how this particular carryover effect of freshwater conditions on marine survival varies by year and rearing type (hatchery or wild), with a larger effect for wild fish. As environmental conditions change, incorporating effects from both freshwater and marine habitats into salmon survival models will become more important, and has the additional benefit of highlighting how management actions that affect arrival timing may improve marine survival.

Keywords: Chinook salmon, hatchery versus wild, marine survival, carryover effects

word count: abstract (208), main (5633), figures (410), tables (220), references (1350)

Introduction

Snake River spring/summer Chinook salmon are an iconic species of the Pacific Northwest. Populations once supported large commercial and recreational fisheries, as well as

subistence for indigenous communities. However, their complex life cycle leaves them vulnerable to the influences of climate and climate change at several life stages [1]. Further, the correlation of environmental conditions across space and time can exacerbate this vulnerability. Indeed, recent research suggests that freshwater effects carry over into the marine realm and may hinder recovery [2–5], but recent applications of generalized linear models to answer this question do not account for the random variability in the carry-over process [6,7]. These simplifications of complex processes in salmon survival models may lead to biases in the parameter estimates and narrow estimates of the standard errors for parameters and derived variables [8], which may compromise our ability to make robust forecasts under future environmental conditions and management scenarios.

Chinook salmon are a semelparous fish with a complex life history, and their survival depends on processes in both freshwater and marine environments over thousands of kilometers [9]. The majority of spring/summer Chinook salmon in the Snake River ESU spend two years in freshwater from the adult spawning migration to the juvenile outmigration and two years in the ocean, with ocean survival showing strong dependence on climatic conditions [10,11]. Data from juvenile Chinook salmon uniquely tagged in the freshwater environment and detected as returning adults suggest that the period when salmon first enter the marine environment is critical to overall marine survival [12–14]. Unfortunately, many of the specific mechanisms of mortality during this period are not well known.

Evaluating drivers of survival for migrating animals is difficult because the interaction between physical processes at local, regional, and basin scales commonly results in correlated conditions across nearby habitat types. In the Columbia River Basin, inter-annual variability in freshwater conditions tends to be correlated with variability in regional marine conditions [7,15],

as both habitats are driven by large-scale atmospheric and oceanographic forces. This correlation has the potential to amplify (or dampen) anomalous conditions in multiple habitats simultaneously, thus complicating our ability to identify causative mechanisms of variability in salmon survival [16].

Moreover, the potential for carryover effects [7] from one environment to another are difficult to incorporate into analyses. Despite evidence showing that the conditions in the freshwater environment play an important role in determining ocean survival by influencing salmon growth rates, condition, and their spatiotemporal proximity to predator populations [6,7,12–14], many questions about carryover effects from the freshwater environment remain. In this effort, we provide a generalized statistical model with which scientists and managers can effectively integrate the complex interacting effects of environmental conditions across multiple habitats. Here, we use a multivariate random effects model to decouple the latent effects of arrival timing from other sources of marine mortality to improve our understanding of the factors affecting the survival of Snake River spring/summer Chinook salmon between 2000 and 2015. The analysis is based on a rich dataset consisting of over 285,000 individually-tagged fish.

Methods

Fish data

We used Passive Integrated Transponder (PIT) data provided by Columbia Basin Research (CBR, cbr.washington.edu) via PIT Tag Information Systems (PTAGIS, www.ptagis.org) to estimate the survival of juvenile salmon. We considered data for all out-migrating yearling spring/summer Chinook salmon tagged in the Snake River Basin detected from 2000 to 2015 at Bonneville Dam -- the furthest downstream dam on the Columbia River.

We marked a fish as having survived the marine stage if it was detected at Bonneville Dam (or farther upstream for less than 2% (5,712 out of 285,600) of fish that were missed at Bonneville Dam) as an adult. The data included i) last detection date at Bonneville Dam as juveniles, ii) rear type (hatchery or wild), and iii) whether the fish was detected in the Columbia River as an adult. We excluded all fish with an unknown rearing type (i.e., hatchery versus wild), geographic regions with fewer than 200 individuals (over the 16 years), those fish released or tagged below the confluence of the Snake and Columbia Rivers, fish that returned to spawn in less than one year, and fish that were barged downstream as juveniles. We excluded fish that passed Bonneville Dam prior to April 9th (day 100) or later than July 8th (day 190); these fish account for <0.14% of the total observations. This resulted in 285,244 individuals for this analysis (Table 1). All PIT-tag files are available on the CBR website (http://www.cbr.washington.edu/dart/cs/data/nmfs_sar/).

Environmental data

Because early ocean experiences are thought to have a large influence on salmon ocean survival [3–5], we focused environmental correlates on marine conditions spanning the winter prior to when fish out-migrated to the fall after outmigration. We obtained these environmental covariate data from a variety of sources (Table 2). Variables represent large-scale oceanographic patterns as well as regional physical metrics. While not all variables have a proposed direct mechanistic relationship with salmon survival, salmon survival has been shown to correlate with many of them [17,18]. Furthermore, we binned all environmental data into three-month averages: these seasonal metrics include Dec-Feb (‘win’), Mar-May (‘spr’), Jun-Aug (‘sum’), and Sep-Nov (‘aut’). These seasonal bins are identified as suffixes on the environmental data

names. For all variables included in our analyses, we tested each of the four seasons, starting with the winter prior to when salmon enter the ocean.

Estimation and data processing

All of the data we used for this analysis are publicly available. We provide a description of the R scripts used to create these environmental data objects from the raw data inputs in supplemental Table.S1. The estimation of the model parameters was done with Template Model Builder (TMB) – a package of C++ libraries that efficiently estimates fixed effects of the model using the AutoDiff libraries and a Laplace approximation to integrate over the random effects.

Models

We used a mixed-effect logistic regression model to predict the SAR for fish of each rear type r (i.e., hatchery versus wild) migrating past Bonneville Dam on calendar day j during year t .

Equation 1.
$$s_{rjt} = \frac{e^{\eta_{rjt}}}{1 + e^{\eta_{rjt}}}$$

Equation 2.
$$\eta_{rjt} = \mu_r + \beta_r x_t + v_{rj} + \omega_{rt} + \varepsilon_{rjt}$$

Where the link function η_{rjt} is a linear combination of fixed and random effects. The fixed-effects consist of the mean survival for rear type r , μ_r , and the vector of marine coefficients for the rear type r , β_r , times the vector of marine covariates in year t , x_t . The random effects include an effect for calendar day j of rear type r , v_{rj} , an effect for year t of rear type r , ω_{rt} , and the interaction between calendar day j and year t for rear type r , ε_{rjt} . A complete list of the subscripts, parameters and data are available in Table 3.

119 The predicted SAR (s_{rjt}) given the observed total cohort of smolts of rear type r that
 120 migrated downstream past Bonneville Dam on day j during year t , n_{rjt} , and the number of fish
 121 from that cohort of rear type r , on day j , and year t that survived, k_{rjt} , is binomially distributed

122 Equation 3. $k_{rjt} \sim \text{Binomial}(n_{rjt}, s_{rjt})$

123 The random effects by rear type r for calendar day and year are treated as auto-regressive
 124 processes with lag 1 (i.e., AR(1)),

125 Equation 4. $v_{rj} \sim \text{Normal}(\rho^{(rj)}v_{rj-1}, (\psi^2)^{(rj)})$

126 Equation 5. $\omega_{rt} \sim \text{Normal}(\rho^{(rt)}\omega_{rt-1}, (\psi^2)^{(rt)})$

127 where, $\rho^{(rj)}$ and $\rho^{(rt)}$ are the correlations for the day and year effects, v_{rj} and ω_{rt} , respectively,
 128 and $(\psi^2)^{(rj)}$ and $(\psi^2)^{(rt)}$ are measures of the variances of the random effects for day and year,
 129 respectively. The random effect for the interaction between day and year was treated as a two-
 130 dimensional auto-regressive process,

131 Equation 6. $\boldsymbol{\varepsilon}_{rt} \sim \text{MVN}(\boldsymbol{\tau}^{(rt)}\boldsymbol{\varepsilon}_{rt-1}, \boldsymbol{\Sigma}^{(r)})$

132 where, $\boldsymbol{\varepsilon}_{rt}$ is a vector of random effects across calendar days for rear type r in year t , $\boldsymbol{\tau}^{(rt)}$ is the
 133 correlation of the vector of day effects between years t and $t-1$, and $\boldsymbol{\Sigma}^{(r)}$ is the covariance matrix
 134 between days within a year for rear type r . The covariance matrix $\boldsymbol{\Sigma}^{(r)}$ is a compact way of
 135 representing the auto-correlation between day effects in the day/year interaction, where the
 136 elements between day j and day $j + \mathbf{D}$ of the covariance matrix are equal to

137 Equation 7. $\Sigma(\varepsilon_{rjt}, \varepsilon_{rj+Dt}) = \frac{(\phi^{(r)})^2}{1 - (\tau^{(rj)})^2} (\tau^{(rj)})^D,$

138 and $(\phi^{(r)})^2$ is variability in the day/year interaction for rear type r.

139 To estimate the fixed and random effects of the model, we use the non-linear
140 optimization libraries in Template Model Builder package [19] built for R [20]. The marginal
141 likelihood of the vector of fixed effects (θ) and the variance parameters (γ) for the random
142 effects (ϵ) given the data (L[Data]) is maximized by integrating across the product of the
143 conditional probability of the data given the fixed and random effects ($\Pr(\theta, \epsilon)$), and the
144 probability of the random effects and the estimated variances ($\Pr(\gamma)$; Thorson & Minto 2014),

145 Equation 8.
$$L[\text{Data}] = \int_{\epsilon} \Pr(\theta, \epsilon) \Pr(\gamma) d\epsilon$$

146 Not all model combinations may be estimable due to the confounding effects among model
147 parameters; in some instances, more than one model parameterization may produce identical fits
148 to the data. In these cases, the Hessian is non-positive definite, and the solution is not unique or
149 estimable. We define a converged model as one with a positive definite hessian and a maximum
150 gradient of 0.001 for the fixed effects. To compare models and select the most parsimonious fit
151 to the data, we used the marginal AIC for the fixed effects (Akaike's information criterion; [21])
152 using the TMBhelper package.

153 Testing all of the thousands of parameter combinations for the 31 marine variables, in
154 addition to the different combinations of random effects, is not reasonable. We therefore
155 restricted the potential models to only those with i) zero, one, or two marine covariates and ii)
156 only two-covariate models where the correlation between covariates was less 0.7. Furthermore,
157 initial analyses indicated that estimating random effects for day, year, and the day/year
158 interaction in a single model produced an over-fit to the data. Models with all three random
159 effects did converge in some instances, but the magnitude of random effects for either the day or

year were so small ($<1e-4$ in most cases) as to be meaningless. Therefore, we restricted our analysis to no more than two random processes for day, year, and the day/year interaction. This resulted in six different random effect models. Finally, to allow for the most flexibility for a given group of fish, we did not combine the hatchery and wild datasets in a multivariate analysis, but rather ran models for each dataset separately. Further research examining the covariance of these two groups could be considered in future analyses.

Model validation

To validate that the parameters for our model (θ and γ) were estimable and unbiased, we conducted simulations in a three by three factorial design - three separate trials for three different simulation experiments (see supplemental Table.S1 for a list of simulation R scripts). For each trial we generated 500 random data sets of the number of wild smolts that survived to adulthood for each day and year in our study. The process for simulating the number of survivors for the k^{th} random draw consisted of: 1) sampling from the hyper-distribution, $\text{Pr}(\gamma_{\text{mle}})$, for the temporal random-effects for trial h and experiment e based on the maximum likelihood estimates (γ_{mle}) for the variances from the model with the lowest AIC (see Table 4, 2) calculate the vector survivals, $\mathbf{s}_{k,h,e}$, based on the random effects $\epsilon_{k,h,e}$, and mle estimates for the mean survival and environmental covariates, θ_{mle} , 3) from the binomial distribution for the observations ($\text{Pr}(\mathbf{n}|\mathbf{s}_{k,h,e})$; see Equation 3), randomly draw the number of fish that survive to adulthood based on the vector sample sizes for each day and year with observations (\mathbf{n}), and the simulated survivals, $\mathbf{s}_{k,h,e}$, and 4) compare the true parameters (θ_{mle} and γ_{mle}) to estimated parameters for the k^{th} simulated data from trial h and experiment e ($\hat{\theta}_{k,h,e}$ and $\hat{\gamma}_{k,h,e}$).

The first experiment set the simulated sample sizes equal to 50%, 100%, and 500% of the observed sample sizes. The second experiment examined the correlation for the day effect by fixing $\rho^{(r=wild,j)}$ to 0.1, 0.5, or 0.9. Similarly, the third experiment examined the effects of the correlations for the day/year interaction by fixing both $\tau^{(r=wild,j)}$ and $\tau^{(r=wild,t)}$ equal to 0.1, 0.5, or 0.9.

Previous models for estimating SAR used generalized linear models (glm) and treated the temporal variability in survival with day, day², and the day/year interaction [6,7] - while Gosselin et al. 2018 did include random variability in the day effect, the process was fixed to being quadratic. We were interested to determine if the previous models could lead to differences in the expectation and standard error for the fixed effects associated with the mean survival and the environmental covariates when compared to our random effects model with AR1 process for day and day/year interactions. Treating our random effects model as the operational model, we simulated 500 data sets for wild fish based on the mle parameters (Table 4), and then compared the fixed effects and standard errors using both the random effects model in TMB and the glm method in R. For the glm model, we used the following log-link function,

$$\text{Equation 9. } \eta_k^{glm} = \mu^{glm} + \beta_{PDO}^{glm} SST_t + \beta_{CUI}^{glm} CUI_t + \beta^{(j,glm)} j + \beta^{(j^2,glm)} j^2 + \beta^{(j*yr,glm)} j * yr.$$

where, for the kth simulation, μ^{glm} is the mean survival, $\beta_{PDO.sum}^{glm}$ and $\beta_{CUI.spr}^{glm}$ are the fixed effects for the environmental covariates, and $\beta^{(j,glm)}$, $\beta^{(j^2,glm)}$, and $\beta^{(j*yr,glm)}$ are fixed effects for the day, day², and day/year interaction, respectively. We selected these environmental covariates because they performed best in our random effects model.

201 Additionally, we used the area under the curve (AUC) statistic based on the receiver
202 operator characteristic (ROC) graphs in the R package pROC [22]. The AUC statistic
203 summarizes the model's ability to discriminate between true positive and false positive rates for
204 a range thresholds. For ecological models, AUC values below 0.7 suggest poor discrimination in
205 the model, while values between 0.7 and 0.8 suggest an acceptable level of discrimination.

206 **Results**

207 *Model fit and best model structure*

208 We found that for wild fish, the models with random effects for day and day/year interactions
209 along with two marine covariates produced the best model fit to the data, and for hatchery fish,
210 models with only day/year interactions and two environmental covariates produced the best fit
211 (Table 4). The top models ($\Delta AIC \leq 4$) for wild fish all assumed random effects for day and
212 day/year interactions, with differences in model fit arising from the combinations of marine
213 covariates (Table 4). A small set of covariates informed the top models for hatchery fish and
214 there was little evidence for an underlying day effect: the only top model for hatchery fish with a
215 day effect had a ΔAIC equal to 4.

216 Comparing the best-fit models for each rearing type, our results suggested that the expected
217 survivals and 95% credible intervals for wild and hatchery fish were 0.009 (0.002, 0.035) and
218 0.008 (0.006, 0.010), respectively (Table 5). The marine covariates that improved the fit of the
219 survival model were different for wild and hatchery fish, but the magnitude of the environmental
220 effects was similar for the two rearing types (Table 5, Figure 1). Spring coastal upwelling index
221 (cui.spr) and summer Pacific decadal oscillation (pdo.sum) provided the best fit to the wild fish
222 data, while summer transport (transport.sum - a measure of the northward transport of water

based on the Sverdrup index) and the summer north Pacific gyre oscillation index (npgo.spr) provided the best fit to the hatchery fish data. The percent change in marine survival as a function of the marine covariates varied between -70% to 150% for wild fish, and -70% and 200% for hatchery fish (Figure 1).

Across all of the top models ($\Delta AIC < 4$), we found differences in the importance of the marine covariates that best explained hatchery and wild survival based on AIC weights (see Figure 2 for calculation). For the top models listed in Table 4, the coastal upwelling index (CUI), Washington coastal and arc sea surface temperatures (ersstWAcost and ersstArc, respectively), and Pacific decadal oscillation (PDO) were important for wild fish (Figure 2), while transport and North Pacific gyre oscillation (NPGO) were most important for hatchery fish. Aside from spring upwelling, covariate indices in summer had stronger weight than other seasons (Figure 2).

For wild fish there was consistently higher survival for the earlier arriving fish - hence, the model with the lowest AIC included the random effect for day (Table 4). The interaction between day and year was important in the best-fit models for both wild and hatchery rearing types (Table 4 and Table 5). Differences in the estimated daily survival rates varied from 0.002 to 0.115 for wild fish, and from 0.003 to 0.06 for hatchery fish (Figure 3). For the day/year effect on the survival of wild fish, there was a strong positive correlation among days within a year ($\tau^{(r=wild,j)} = 0.932$), and negative correlation among days across years ($\tau^{(r=wild,t)} = -0.489$) (Table 5). The random deviation of the day/year interaction for hatchery fish showed a high degree of correlation among days within a year ($\tau^{(r=hatchery,j)} = 0.955$) and a weak negative correlation among days across years ($\tau^{(r=hatchery,t)} = -0.067$). The standard deviation of the day/year interactions was similar for hatchery fish ($(\psi)^{(hatchery)} = 0.611$) and wild fish ($(\psi)^{(wild)} = 0.58$).

To illustrate the effect of arrival timing for wild and hatchery fish, we compared the top model for each rearing type that included the random effects for both the day and the day/year interactions. For wild fish, this was the model with the lowest AIC, and for hatchery fish, this was a model with identical marine covariates to the best-fit model but with daily random effects ($\Delta\text{AIC} = 3.9996$; Table 4). The day effect was highest for wild fish passing Bonneville Dam around May 3rd, followed by decreasing survival throughout the remainder of the smolt migration (Figure 4a). By comparison, the model of hatchery fish that included both day and day/year interaction showed no real difference in smolt survival for the day effect (Figure 4a), despite relatively similar mean arrival timing past Bonneville Dam (Figure 4b). The lack of a day effect for hatchery fish is supported by the low estimates for the correlation coefficient and variability in their day effect ($\rho^{\text{hatchery},j} = 0.05$ and $\phi^{\text{hatchery},j} = 0.134$). Conversely, the wild fish had higher correlation and variability ($\rho^{\text{wild},j} = 0.986$ and $\phi^{\text{wild},j} = 0.793$, respectively) which suggests that the day effect “wanders” more for wild fish.

While none of the top models included a random deviate for year, we predicted the annual survival by aggregating the daily survival estimates weighted by the total number of hatchery and wild fish that arrived each day. The observed annual survival estimates were similar to the model predictions and, with the exception of wild fish in 2003, the observations fell within the 95% credible interval (Figure 5). Both the predicted and observed annual survivals for hatchery and wild fish showed an alternating pattern of increases and decreases, which was evident by the previously described negative correlations in the year dimension for the day/year interaction (Table 5).

Model validation

The AUC statistics for the hatchery and wild fish models were equal to 0.69 and 0.76, respectively. This indicates that the ability for the hatchery fish model to discriminate between true positives and false positives was barely below the acceptable threshold of 0.7, while the wild fish model was above it. Visual inspection of the simulation experiments suggests that the estimation model provided unbiased estimates (i.e., the center of mass of the violins is near zero) for the fixed effects in the TMB model (Figure S1). Across all experiments, some random draws led to negative biases for the standard deviations of the day effect (ψ_j). These biases were usually associated with random draws with low numbers of surviving fish and little autocorrelation in the day effect. As sample sizes increased, the precision increased (i.e., the violins get vertically compressed) for the correlations and standard deviations that describe the random processes for day and day/year. Experiments examining the magnitude of the correlation for both the day ($\rho^{(j)}$) and day/year ($\tau^{(j)}$ and $\tau^{(t)}$) effects resulted in no bias in the other fixed effects (i.e., average deviation was zero). Additionally, as the magnitude of the correlations for the random processes increased, the precision increased for the estimated correlations, decreased for the mean survival (μ), and remained unchanged for the marine covariates ($\beta_{\text{PDO.sum}}$ and $\beta_{\text{CUI.spr}}$).

When we compared the performance of the generalized linear model (glm) with fixed-effects for day, day², and day/year interaction using the glm function in R with the random effects model in TMB, we found little difference in the bias and precision of the mean survival (μ) and the marine covariates ($\beta_{\text{PDO.sum}}$ and $\beta_{\text{CUI.spr}}$) (Figure 6, upper panel). Examining a single realization for the simulated data, the survival estimates from glm model (Figure S2, red line) and TMB model (blue line) were similar. However, the random processes in the TMB model provided more flexibility to capture the daily variability in survival within a year. While there were almost no differences in the bias or precision for the mean survival and environmental

covariates, the standard errors for those fixed-effects in the TMB model were between 65%-70% higher than the glm model (Figure 6, lower panel).

Discussion

We found that for Snake River spring/summer Chinook salmon, survival in the ocean was strongly related to several indicators of ocean conditions but also to when individuals arrived at the estuary. Arrival timing is the culmination of processes that occur in the freshwater, so we have established strong linkages between freshwater conditions and ocean conditions. Our modeling framework allows for historical trends but also has the flexibility to forecast trends into the future. Perhaps counterintuitively, increasing the flexibility of the model and allowing more of the uncertainty to be explained by these temporal processes led to increased uncertainty in the mean survival and environmental covariates (Figure 6). Thus, this research highlights that conclusions about the uncertainty in the survival estimates must also reflect the uncertainty in the processes that are believed to affect survival (i.e., timing). Additionally, our generalized approach for integrating random variability into an SAR model can easily be expanded from the AR1 “lattice” for the day/year interaction to higher dimensional interactions that include biological forces such as size and weight, or environmental forces such stream temperature. Because these forces are associated with “levers” that managers of freshwater systems can manipulate – as opposed to climate conditions, quantifying the effect of these interactions on the uncertainty in survival is critically important for evaluating future management scenarios.

Hatchery-wild comparisons

Different rearing types of spring/summer Chinook salmon exhibit differences in SAR within and between years. There are expected differences between fish reared in a hatchery and fish exposed to natural conditions in the wild, including size, condition, risk aversion, arrival

timing, parasite load, and numerous other factors. We clearly documented the effect of arrival timing on marine survival was not consistent between fish of different rearing types, and we described two primary differences in timing and marine survival. First, the arrival timing distribution for juvenile salmon differs between the hatchery and wild fish, with hatchery fish clumped in a narrow window, mostly completed by early June. In contrast, wild fish start to arrive earlier and the distribution has a long tail, with some fish not passing Bonneville Dam until mid-July. Second, on average across years, survival peaks for wild fish migrating early and then declines throughout the remainder of the migration, whereas hatchery fish, on average, show no consistent pattern in survival across years based on arrival timing.

There are multiple reasons why wild fish may be more sensitive to arrival timing than hatchery fish, though much of this is speculation. Perhaps the most likely cause is the difference in size between the two groups. If early marine survival is size-dependent, which has been shown for other salmon stocks [13,23–25], the larger size of hatchery fish could afford them some level of independence from predators. Additionally, large subsidies of hatchery smolts may increase the density of the predator communities, and these predators may differentially select for wild fish because they are smaller and more available once the pulse of hatchery fish has passed [12].

Arrival timing

A key component of this model is the inclusion of arrival timing to the marine environment. Gosselin et al. [7] showed that management practices in freshwater can have large impacts on marine survival via carryover effects, which can materialize in the form of altered fish size or timing at out-migration. Although size-dependent mortality is important, we focused on the impacts of timing for this effort. Arrival timing has been shown to be an important

catalyst for carryover effects [7] and these data are quite readily available, as each fish detected at Bonneville Dam has its own time stamp. However, there is a large amount of variability in arrival timing, and managers of wild salmon populations have few levers to manipulate the environmental experiences that may influence marine survival. To the extent that the freshwater environment influences salmon behavior, performance, growth, and survival in the marine environment, these influences should be incorporated into modeling efforts aimed at understanding salmon marine ecology. Freshwater conditions affecting arrival timing (e.g., flow and temperature) are likely to be correlated with conditions in the marine environment [26], and phenological variability in the marine ecosystem is driven by atmospheric and oceanographic processes with substantial inter-annual variability [27]. For example, wind-driven ocean currents transition from south to north each spring, initiating a strong upwelling of deeper ocean water. The nutrients in this upwelled water can spawn or feed a spring phytoplankton bloom [28]. Moreover, the newly transitioned currents can bring species of zooplankton such as copepods that are high in fatty acids [29,30], further enriching the production at lower trophic levels. Salmon eventually benefit from these dynamics, but the timing and magnitude of local production varies from year to year. Although salmon have evolved to optimize arrival timing on average [31], the broad distribution of arrival timing may be a bet-hedging strategy [32,33] to ensure some fish arrive at the ocean when conditions are optimal. If future freshwater management practices significantly alter the mean arrival timing or the variability in timing, this could have important, and perhaps unforeseen, effects on marine survival. Similarly, if climate changes in either the freshwater or marine environment result in a mismatch between salmon arrival timing and optimal arrival timing, marine survival will be impacted. These interactions

are a clear demonstration of the importance of carryover effects and a direct link between salmon survival and management decisions that may affect arrival timing [34].

Marine covariates

The top models describing Chinook salmon marine survival included three categories of environmental covariates for wild fish (i.e., basin-scale sea surface temperature ('ersstArc' and PDO), a local sea surface temperature ('ersstWA'), and a regional spring upwelling variable ('cui')), and three categories of environmental covariates for hatchery fish (i.e., a measure of alongshore flow ('transport'), ocean circulation ('NPGO'), and sea surface temperature ('errstArc')). For each rearing type, there are logical links between the metric and multiple oceanographic or ecosystem processes that could influence salmon growth and survival. However, most of these links are indirect and rely on other oceanographic factors. For example, local sea surface temperature can influence growth rates directly [35], but a more likely influence on salmon survival involves production at lower trophic levels and temperature-dependent distribution of prey and predator species [36].

In this effort, we intentionally restricted our potential ocean covariates to publicly-available (and mostly physical) variables. These variables do not necessarily directly relate to the ecosystem processes that determine salmon survival, but rather represent correlations with these processes. Some biological time series that more directly characterize ecosystem processes such as trophic dynamics are available, but only for recent years (e.g., stoplight chart for ocean survival estimates, <http://www.nwfsc.noaa.gov/oceanconditions>). For other research goals, such as near-term forecasting, these more direct metrics may be more appropriate. Indeed, as more biological data are collected, reliance on correlations should be reduced [37] and the use of mechanistic ecosystem models will become more important [38,39].

Model Fit

Comparing the residual deviance ratio, defined as the fit of a particular model relative to the model where each data point has its own parameter, the fixed effects models that included only marine covariates had ratios equal to 0.077 and 0.167 for wild and hatchery fish, respectively. When we removed the marine covariates and included a day/year interaction, the ratios increased to 0.197 and 0.346, respectively. Finally, the ratios increased to 0.208 and 0.350, respectively, for the model that included marine covariates and random effects for day and the day/year interaction (Table 6). The small differences in the ratios between the random effects models with and without marine covariates does not imply that marine conditions do not affect Chinook salmon survival. In fact, as shown by the estimated magnitude of the deviates in Figure 1, the marine covariates were correlated with large differences in marine survival. However, rather than a uniform response of all fish to the marine conditions in a particular year, our model demonstrates that the timing of when the juvenile salmon encounter the marine conditions appears to explain more of the data (Table 6). The mechanism that is driving this differential survival across days and years remains a critical knowledge gap and a focus of future salmon modeling.

Combining impacts from multiple environments has been applied in several past efforts to model Snake River spring/summer Chinook marine survival. The day effect was described by Scheuerell et al. [6] using a quadratic effect for day in a logistic regression model and showed that earlier fish tend to have higher survival, but this shifted somewhat from year to year. Holsman et al. [40] also use a logistic regression for this ESU and characterized the impacts of predators, prey, flow, and the temperature difference between the Columbia River and the nearshore ocean; however, they did not include a day effect in their model. Similarly, Haseker et

al. [41] demonstrated the importance of river flow (the proportion of water spilled over dams and migration rate), in modeling marine survival for this ESU, but included a linear effect of day. Miller et al. [42] used a logistic regression to show that the size at out-migration was not as important as the size at marine capture (after fish had been in the ocean for weeks to months), suggesting that marine growth is highly influential in setting mortality rates. Finally, Gosselin et al. [7] used a mixed effects regression to describe carryover effects from the freshwater environment, with particular emphasis on transportation impacts on hatchery and wild fish, but constrained the underlying process for the day effect to be quadratic. Our current model design represents a compromise between model complexity, realism, and the clear need to address the interactions between freshwater impacts and the marine ecosystem. Rather than treating the effect of timing on survival as a fixed effect described by a linear or quadratic relationship, our model accounts for the heterogeneity in the survival processes by treating the effect of timing as random process.

Using the model prospectively

We have demonstrated that our model is powerful for detecting effects in both marine and freshwater environments from historical data. However, we designed the model such that it can also be used for population viability modeling. To do this, the ocean survival model is incorporated into a stochastic age-based life cycle model (e.g., Zabel et al. 2006). This approach has been adopted by NOAA Fisheries to examine the effects of climate and climate change on salmon population viability [43]. The fact that several of the most important ocean indicators (e.g., SST) are amenable to forecasting under climate change scenarios allows for an important examination of how Snake River spring/summer will respond to future climate conditions.

427 *Caveats*

428 We included arrival timing, but did not include other attributes such as fish size, which is
429 known to have important impacts on trophic interactions and size-dependent survival
430 [13,24,25,44]. Miller et al. [42] showed that Snake River spring/summer Chinook marine
431 survival was more related to size after some period of ocean residence than size at out-migration,
432 but did not rule out the possibility that some level of size-dependent mortality did not already
433 occur. Further research to extend this model is necessary to fully understand how other fish
434 attributes such as size in the freshwater environment are likely to affect marine survival.
435 Fortunately, given the flexibility of the multivariate framework, such analyses are possible with
436 the availability of additional data. Additionally, maturation schedules, the fraction of a salmon
437 maturing and returning to spawn at different ages, are also size-dependent – larger and faster
438 growing fish tend to mature earlier [2]. Recent spawner-recruit analysis suggests that climate
439 conditions affect both the maturation schedule and the survival of some stocks of salmon [45];
440 however, timing and size were not a part of these models. Future iterations of our model could
441 examine the effects of size and maturation simultaneously, with the goal of understanding how
442 management actions in freshwater environment affect size, maturation, and ultimately, survival.

443 We view our model as a robust approach for integrating the freshwater and marine effects
444 in a single estimation model. By partitioning the different sources of uncertainty between the
445 observation model (binomial likelihood) and process models (random effects for day, year, and
446 day/year interactions) we provide a more accurate estimate of the uncertainty and relative
447 importance of the fixed effects associated with the marine covariates relative to the random
448 deviations in survival associated with differences in arrival timing between years. While our
449 model was restricted to examining the two-dimensional interaction between day and year, this

model can quickly be scaled-up to higher-dimensional questions related to the interaction between day, year, size, and maturation.

Acknowledgements

We would like to thank Jeff Jorgenson and David Huff for their reviews of previous versions of this manuscript, Susan Iltus for helping us to compile PIT tag data for this analysis, and Jennifer Gosselin for listening to our initial formulations of the model.

References

1. Crozier LG, McClure MM, Beechie T, Bograd SJ, Boughton DA, Carr M, et al. Climate vulnerability assessment for Pacific salmon and steelhead in the California Current Large Marine Ecosystem. *PloS one*. 2019;14.
2. Scheuerell MD, Williams JG. Forecasting climate-induced changes in the survival of Snake River spring/summer Chinook salmon (*Oncorhynchus tshawytscha*). *Fisheries Oceanography*. 2005;14: 448–457.
3. Kilduff DP, Botsford LW, Teo SL. Spatial and temporal covariability in early ocean survival of Chinook salmon (*Oncorhynchus tshawytscha*) along the west coast of North America. *ICES Journal of Marine Science*. 2014;71: 1671–1682.
4. Woodson CB, Litvin SY. Ocean fronts drive marine fishery production and biogeochemical cycling. *Proceedings of the National Academy of Sciences*. 2015;112: 1710–1715.
5. Wells BK, Santora JA, Schroeder ID, Mantua N, Sydeman WJ, Huff DD, et al. Marine ecosystem perspectives on Chinook salmon recruitment: a synthesis of empirical and modeling studies from a California upwelling system. *Marine Ecology Progress Series*. 2016;552: 271–284.

- 472 6. Scheuerell MD, Zabel RW, Sandford BP. Relating juvenile migration timing and survival
473 to adulthood in two species of threatened Pacific salmon (*Oncorhynchus* spp.). *Journal of*
474 *Applied Ecology*. 2009;46: 983–990.
- 475 7. Gosselin JL, Zabel RW, Anderson JJ, Faulkner JR, Baptista AM, Sandford BP.
476 Conservation planning for freshwater–marine carryover effects on Chinook salmon
477 survival. *Ecology and evolution*. 2018;8: 319–332.
- 478 8. Thorson JT, Minto C. Mixed effects: a unifying framework for statistical modelling in
479 fisheries biology. *ICES Journal of Marine Science*. 2015;72: 1245–1256.
- 480 9. Quinn TP. The behavior and ecology of Pacific salmon and trout. University of Washington
481 press; 2018.
- 482 10. Mantua NJ, Hare S. Pacific-Decadal Oscillation (PDO). *Encyclopedia of global*
483 *environmental change*. 2002;1: 592–594.
- 484 11. Zabel RW, Scheuerell MD, McCLURE MM, Williams JG. The interplay between climate
485 variability and density dependence in the population viability of Chinook salmon.
486 *Conservation Biology*. 2006;20: 190–200.
- 487 12. Beamish RJ, Thomson BL, McFarlane GA. Spiny dogfish predation on chinook and coho
488 salmon and the potential effects on hatchery-produced salmon. *Transactions of the*
489 *American Fisheries Society*. 1992;121: 444–455.

- 490 13. Duffy EJ, Beauchamp DA. Rapid growth in the early marine period improves the marine
491 survival of Chinook salmon (*Oncorhynchus tshawytscha*) in Puget Sound, Washington.
492 Canadian Journal of Fisheries and Aquatic Sciences. 2011;68: 232–240.
- 493 14. Chasco BE, Kaplan IC, Thomas AC, Acevedo-Gutiérrez A, Noren DP, Ford MJ, et al.
494 Competing tradeoffs between increasing marine mammal predation and fisheries harvest of
495 Chinook salmon. Scientific Reports. 2017;7: 1–14.
- 496 15. Hodgson S, Quinn TP. The timing of adult sockeye salmon migration into fresh water:
497 adaptations by populations to prevailing thermal regimes. Canadian Journal of Zoology.
498 2002;80: 542–555.
- 499 16. Litzow MA, Ciannelli L, Puerta P, Wettstein JJ, Rykaczewski RR, Opiekun M. Non-
500 stationary climate–salmon relationships in the Gulf of Alaska. Proceedings of the Royal
501 Society B. 2018;285: 20181855.
- 502 17. Burke BJ, Peterson WT, Beckman BR, Morgan C, Daly EA, Litz M. Multivariate models of
503 adult Pacific salmon returns. PloS one. 2013;8.
- 504 18. Peterson WT, Fisher JL, Peterson JO, Morgan CA, Burke BJ, Fresh KL. Applied fisheries
505 oceanography: Ecosystem indicators of ocean conditions inform fisheries management in
506 the California Current. Oceanography. 2014;27: 80–89.
- 507 19. Kristensen K, Nielsen A, Berg CW, Skaug H, Bell B. TMB: automatic differentiation and
508 Laplace approximation. arXiv preprint arXiv:150900660. 2015.
- 509 20. Team RC. R: A language and environment for statistical computing. 2013.

- 510 21. Akaike H. A new look at the statistical model identification. IEEE transactions on
511 automatic control. 1974;19: 716–723.
- 512 22. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. Package ‘pROC.’
513 2012-09-10 09: 34; 2020.
- 514 23. Henderson MA, Cass AJ. Effect of smolt size on smolt-to-adult survival for Chilko Lake
515 sockeye salmon (*Oncorhynchus nerka*). Canadian Journal of Fisheries and Aquatic
516 Sciences. 1991;48: 988–994.
- 517 24. Beamish RJ, Mahnken C. A critical size and period hypothesis to explain natural regulation
518 of salmon abundance and the linkage to climate and climate change. Progress in
519 Oceanography. 2001;49: 423–437.
- 520 25. Woodson LE, Wells BK, Weber PK, MacFarlane RB, Whitman GE, Johnson RC. Size,
521 growth, and origin-dependent mortality of juvenile Chinook salmon *Oncorhynchus*
522 *tshawytscha* during early ocean residence. Marine Ecology Progress Series. 2013;487: 163–
523 175.
- 524 26. Keefer ML, Peery CA, Caudill CC. Migration timing of Columbia River spring Chinook
525 salmon: effects of temperature, river discharge, and ocean environment. Transactions of the
526 American Fisheries Society. 2008;137: 1120–1133.
- 527 27. Mantua NJ, Hare SR, Zhang Y, Wallace JM, Francis RC. A Pacific interdecadal climate
528 oscillation with impacts on salmon production. Bulletin of the American Meteorological
529 Society. 1997;78: 1069–1080.

- 530 28. Du X, Peterson WT. Seasonal cycle of phytoplankton community composition in the
531 coastal upwelling system off central Oregon in 2009. *Estuaries and coasts*. 2014;37: 299–
532 311.
- 533 29. Hooff RC, Peterson WT. Copepod biodiversity as an indicator of changes in ocean and
534 climate conditions of the northern California current ecosystem. *Limnology and*
535 *Oceanography*. 2006;51: 2607–2620.
- 536 30. Keister JE, Di Lorenzo E, Morgan CA, Combes V, Peterson WT. Zooplankton species
537 composition is linked to ocean transport in the Northern California Current. *Global Change*
538 *Biology*. 2011;17: 2498–2511.
- 539 31. Spence BC, Dick EJ. Geographic variation in environmental factors regulating outmigration
540 timing of coho salmon (*Oncorhynchus kisutch*) smolts. *Canadian journal of fisheries and*
541 *aquatic sciences*. 2014;71: 56–69.
- 542 32. Schindler DE, Hilborn R, Chasco B, Boatright CP, Quinn TP, Rogers LA, et al. Population
543 diversity and the portfolio effect in an exploited species. *Nature*. 2010;465: 609–612.
- 544 33. Griffiths JR, Schindler DE, Armstrong JB, Scheuerell MD, Whited DC, Clark RA, et al.
545 Performance of salmon fishery portfolios across western North America. *Journal of*
546 *Applied Ecology*. 2014;51: 1554–1563.
- 547 34. Crozier LG, Hendry AP, Lawson PW, Quinn TP, Mantua NJ, Battin J, et al. Potential
548 responses to climate change in organisms with complex life histories: evolution and
549 plasticity in Pacific salmon. *Evolutionary Applications*. 2008;1: 252–270.

- 550 35. Wells BK, Grimes CB, Waldvogel JB. Quantifying the effects of wind, upwelling, curl, sea
551 surface temperature and sea level height on growth and maturation of a California Chinook
552 salmon (*Oncorhynchus tshawytscha*) population. *Fisheries Oceanography*. 2007;16: 363–
553 382.
- 554 36. Wells PM, Baverstock J, Clark SJ, Jiggins FM, Roy HE, Pell JK. Determining the effects of
555 life stage, shared prey density and host plant on intraguild predation of a native lacewing
556 (*Chrysoperla carnea*) by an invasive coccinellid (*Harmonia axyridis*). *Biocontrol*. 2017;62:
557 373–384.
- 558 37. Litzow MA, Hunsicker ME, Bond NA, Burke BJ, Cunningham CJ, Gosselin JL, et al. The
559 changing physical and ecological meanings of North Pacific Ocean climate indices.
560 *Proceedings of the National Academy of Sciences*. 2020;117: 7665–7671.
- 561 38. Hollowed AB, Bax N, Beamish R, Collie J, Fogarty M, Livingston P, et al. Are
562 multispecies models an improvement on single-species models for measuring fishing
563 impacts on marine ecosystems? *ICES Journal of Marine Science*. 2000;57: 707–719.
- 564 39. Fulton EA, Link JS, Kaplan IC, Savina-Rolland M, Johnson P, Ainsworth C, et al. Lessons
565 in modelling and management of marine ecosystems: the Atlantis experience. *Fish and*
566 *Fisheries*. 2011;12: 171–188.
- 567 40. Holsman KK, Scheuerell MD, Buhle E, Emmett R. Interacting effects of translocation,
568 artificial propagation, and environmental conditions on the marine survival of Chinook
569 Salmon from the Columbia River, Washington, USA. *Conservation Biology*. 2012;26: 912–
570 922.

41. Haeseker SL, McCann JA, Tuomikoski J, Chockley B. Assessing freshwater and marine environmental influences on life-stage-specific survival rates of Snake River spring–summer Chinook salmon and steelhead. *Transactions of the American Fisheries Society*. 2012;141: 121–138.
42. Miller JA, Teel DJ, Peterson WT, Baptista AM. Assessing the relative importance of local and regional processes on the survival of a threatened salmon population. *PLoS One*. 2014;9.
43. NMFS M. Endangered Species Act (ESA) Section 7 (a)(2) Biological Opinion and Magnuson-Stevens Fishery Conservation and Management Act Essential Fish Habitat (EFH) Consultation. 2019.
44. Roby DD, Lyons DE, Craig DP, Collis K, Visser GH. Quantifying the effect of predators on endangered species using a bioenergetics approach: Caspian terns and juvenile salmonids in the Columbia River estuary. *Canadian Journal of Zoology*. 2003;81: 250–265.
45. Scheuerell MD, Ruff CP, Anderson JH, Beamer EM. An integrated population model for estimating the relative effects of natural and anthropogenic factors on a threatened population of Pacific trout. *bioRxiv*. 2019; 734996.

588 **Figure captions**

589 Figure 1. Environmental effects on survival of wild (upper panel) and hatchery (lower panel)
590 spring/summer Chinook salmon based on the model fit to the observed data as selected by AIC
591 (see Table 4 for summary of best fit models).

592 Figure 2. Relative importance of the different marine covariates for predicting the marine
593 survival of hatchery (left column) and wild (right column) spring/summer Chinook salmon,
594 where the aggregated weight of a covariate c is equal to the sum of the AIC weights for all m
595 models containing covariate c , divided by the total weight across all m models, $\frac{\sum_{c \in m} e^{(\Delta AIC_m)}}{\sum_m e^{(\Delta AIC_m)}}$.
596 The “blank” environmental variable is for models with no environmental predictors.

597 Figure 3. Split violin plot comparing the percent difference (upper panel) between the estimated
598 and true parameter values and the standard errors for the fixed effects μ , $\beta_{CUI.spr}$, and $\beta_{PDO.sum}$
599 (lower panel) for the mixed-effect model with AR1 processes for day and day/year interaction
600 (TMB; blue violins) and the fixed-effect model for day, day², and day/year interaction (glm; red
601 violins) fit to simulated data for wild spring/summer Chinook salmon. Simulated the data were
602 generated from the model with the lowest AIC for wild fish (see Table 4). Horizontal lines
603 represent median values for the violins and the horizontal blue line in panel (A) represents zero
604 percent difference between the estimated and true parameter values.

605 Figure 4. The observed (dots), and maximum likelihood estimates (line) with 95% credible
606 intervals (ribbons) for the marine survival wild (blue) and hatchery (red) origin Spring/Summer
607 Chinook salmon past Bonneville dam from 2000 to 2015. Each point represents the mean
608 survival of all fish detected at Bonneville Dam on a particular day and year. Annual samples
609 sizes of the survivors and total PIT tagged hatchery (H) and wild (W) for are shown in each

610 panel. To maintain the readability of individual panels, mean observed survivals greater than 0.2
611 are not plotted.

612 Figure 5. Predicted smolt-to-adult survival by day for hatchery (red) and wild (blue)
613 spring/summer Chinook salmon (upper panel) for best fit models for each rear type that include
614 both day and day/year interactions (see Table 4). Lines represent expected survivals and shaded
615 regions represent 95% credible intervals. Average daily proportion (across all years) of smolts
616 arriving to and migrating past Bonneville Dam (2000 to 2015) (lower panel).

617 Figure 6. Observed (points) and estimated (line) annual survival with 95% credible intervals
618 (polygons) for hatchery (red) and wild (blue) spring/summer Chinook salmon from 2000 through
619 2015.

620

621 **Table captions**

622 Table 1. Sample size of hatchery and wild juvenile Chinook salmon from different watersheds
623 within the Snake River Basin.

624

625 Table 2. Description of the environmental variable names, the years of available data, and the
626 website location of the data.

627

628 Table 3. List of data types, subscripts, parameters (i.e., fixed effects), and random effects used to
629 model the smolt to adult survival of spring/summer Chinook salmon originating from the Snake
630 River Basin.

631

632 Table 4. Top models for each origin type based on the random effects and number of marine
633 covariates. The gradient is a measure of the likelihood surface for the maximum likelihood
634 estimate, PD Hessian stands for positive definite Hessian, and N.E. stands for not estimated.

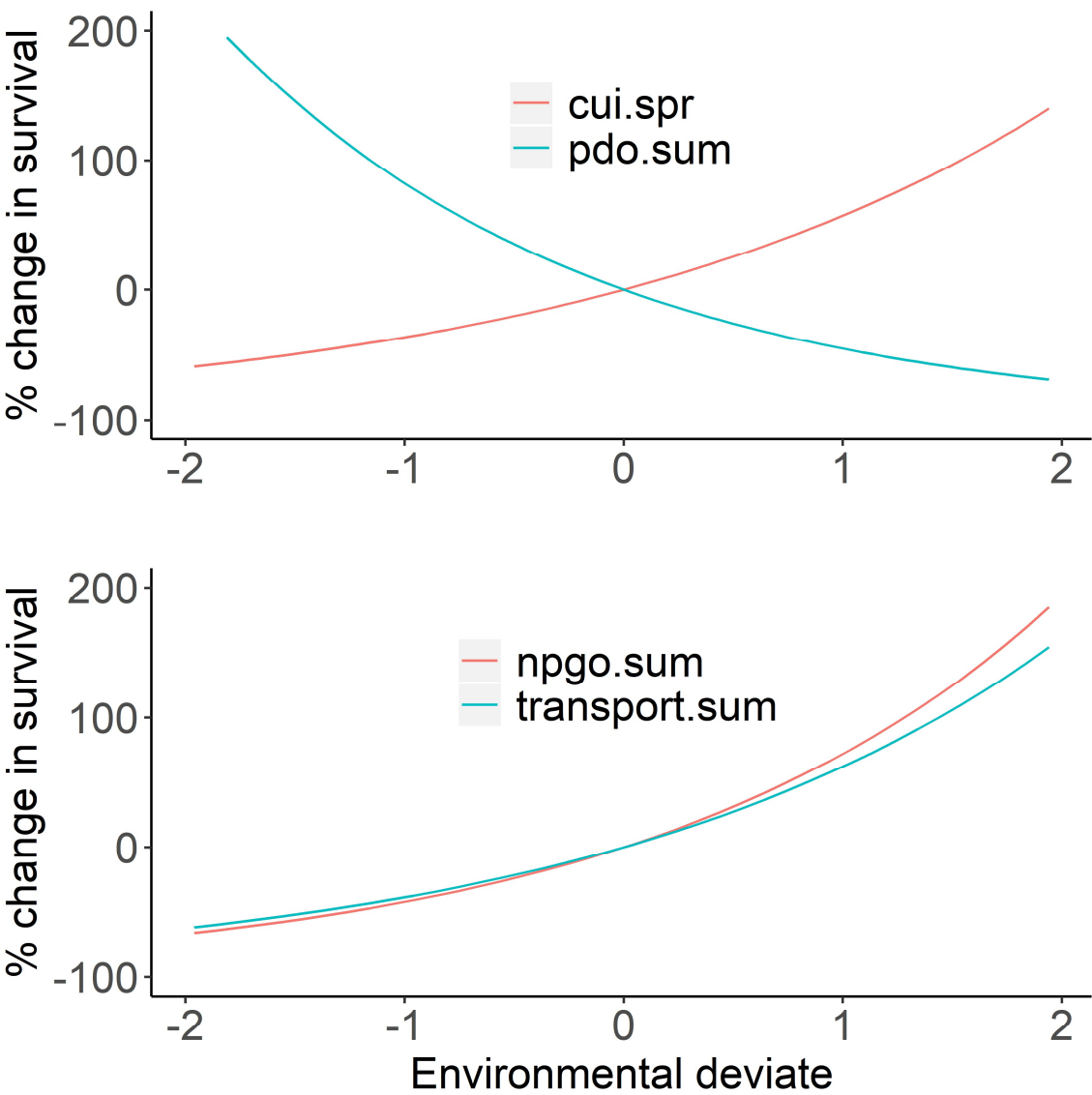
635

636 Table 5. Estimated fixed effects for the model with lowest marginal AIC values for wild
637 spring/summer Chinook salmon, and the models with the lowest marginal AIC for hatchery fish
638 with and without a random day effect included (N.E. means not estimated).

639 Table 6. Deviance ratios $\left(\text{i. e.}, 1 - \frac{D_m}{D_0} \right)$ for different fixed and mixed-effects models for hatchery
640 and wild fish, where D_m is -2 times the log-likelihood of model m and D_0 is -2 times the log-

641 likelihood of the null model. The deviance ratio explains how close the model is to the best
642 model that fits the data exactly.

643



645

646 Figure 1. Environmental effects on survival of wild (upper panel) and hatchery (lower panel)

647 spring/summer Chinook salmon based on the model fit to the observed data as selected by AIC

648 (see Table 4 for summary of best fit models).

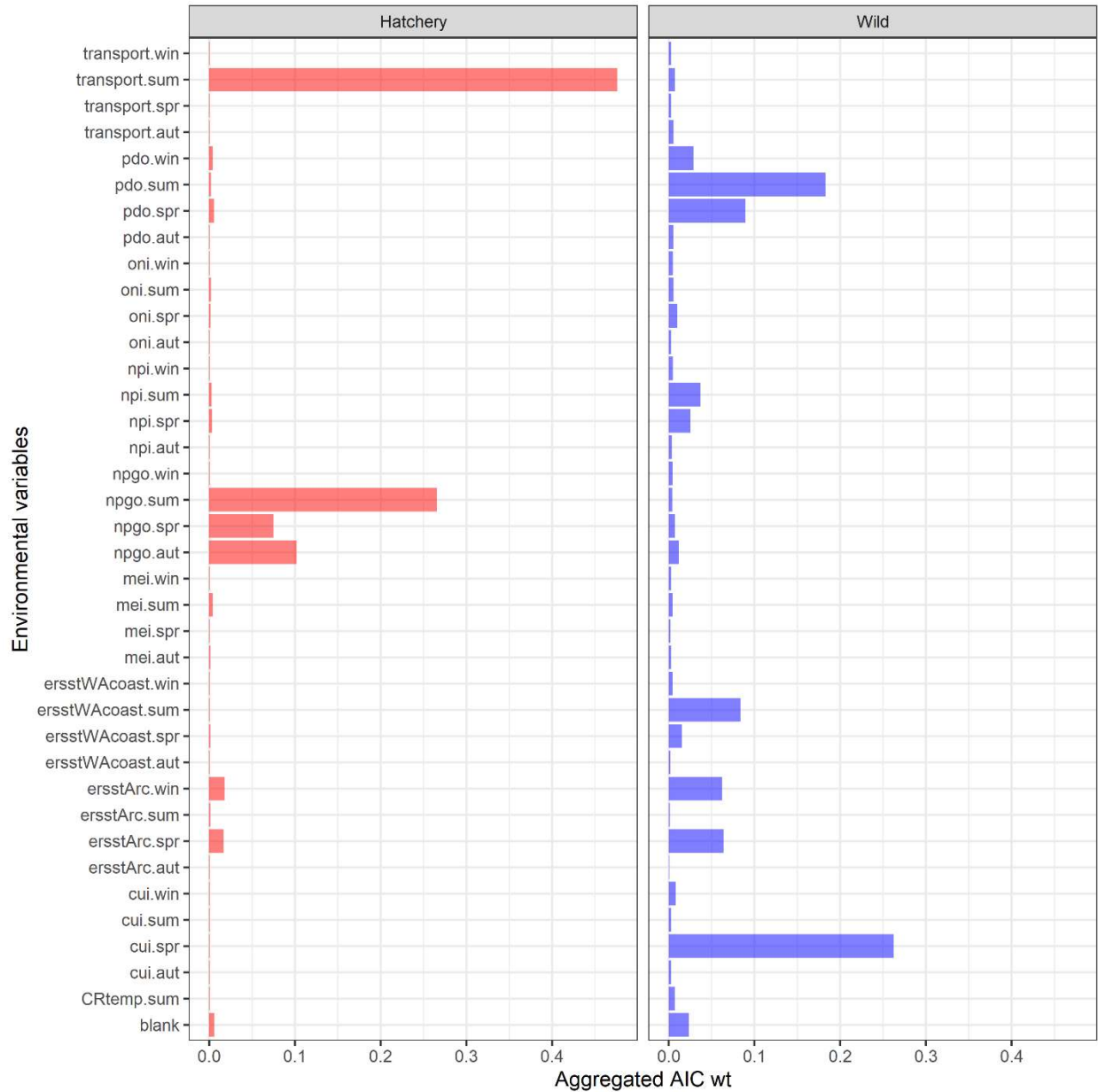
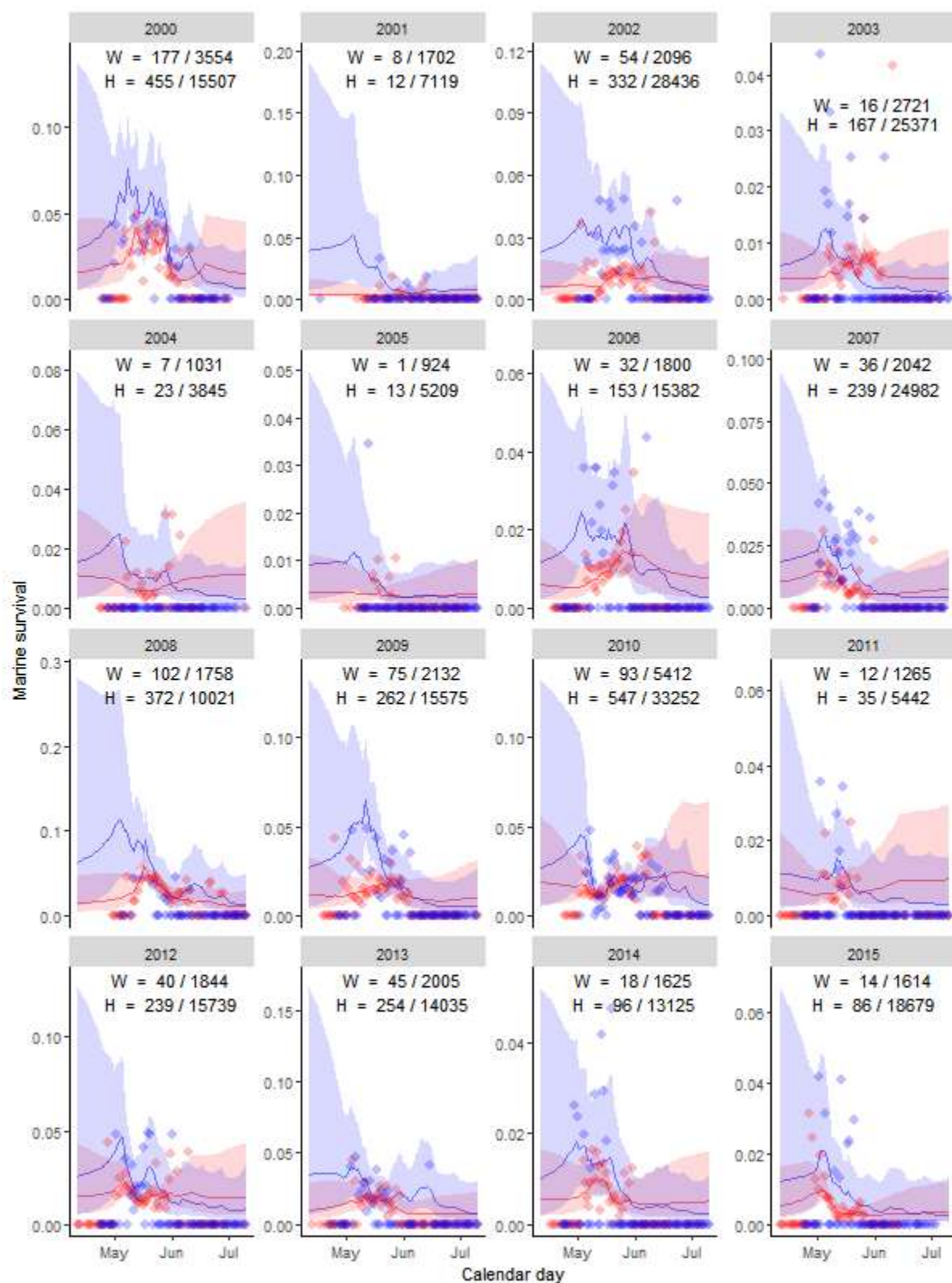


Figure 2. Relative importance of the different marine covariates for predicting the marine survival of hatchery (left column) and wild (right column) spring/summer Chinook salmon, where the aggregated weight of a covariate c is equal to the sum of the AIC weights for all m models containing covariate c , divided by the total weight across all m models, $\frac{\sum_{c \in m} e^{(\Delta AIC_m)}}{\sum_m e^{(\Delta AIC_m)}}$. The “blank” environmental variable is for models with no environmental predictors.



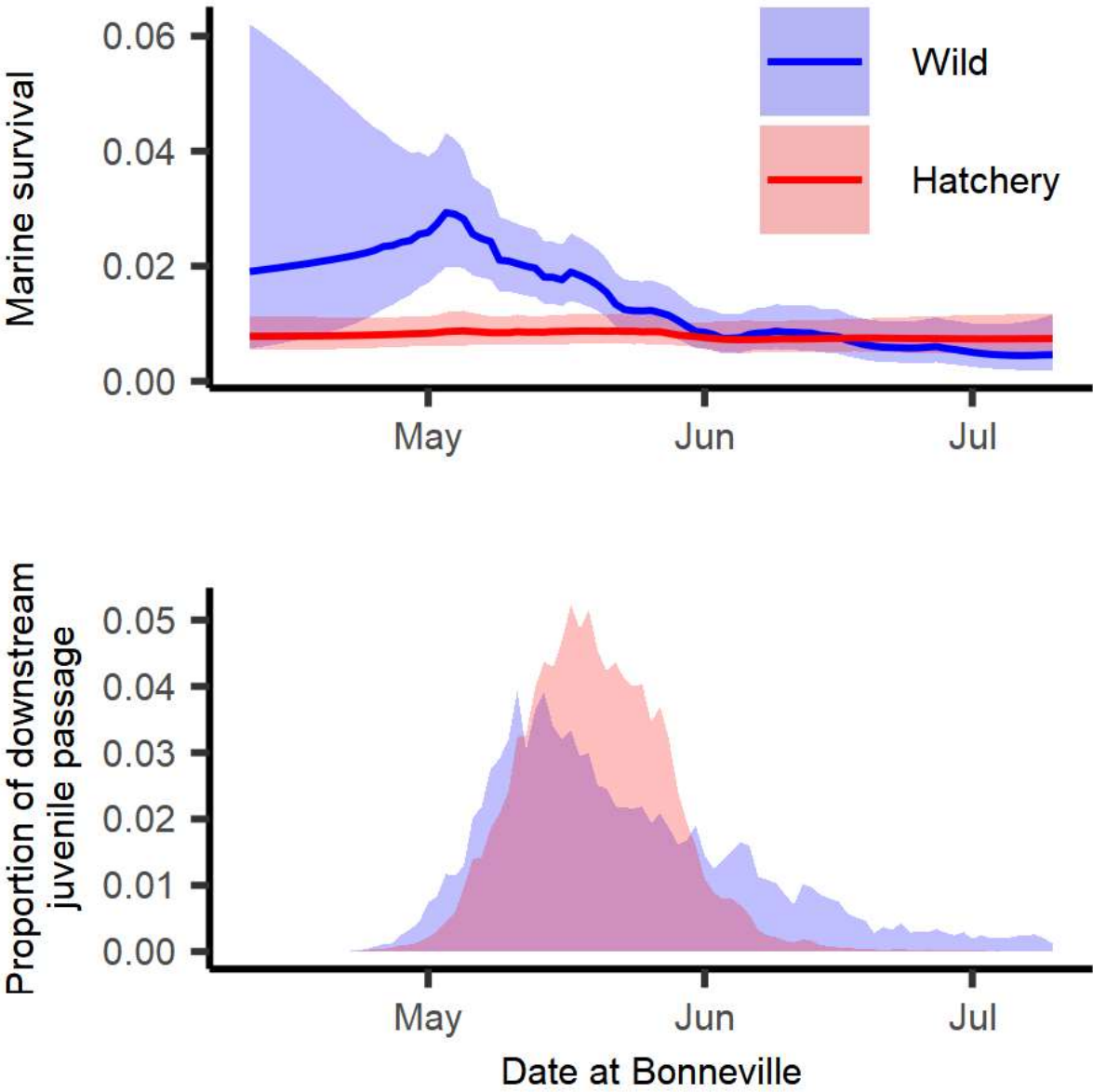
658

659

660 Figure 3. The observed (dots), and maximum likelihood estimates (line) with 95% credible
661 intervals (ribbons) for the marine survival wild (blue) and hatchery (red) origin Spring/Summer
662 Chinook salmon past Bonneville dam from 2000 to 2015. Each point represents the mean
663 survival of all fish detected at Bonneville Dam on a particular day and year. Annual samples
664 sizes of the survivors and total PIT tagged hatchery (H) and wild (W) for are shown in each
665 panel. To maintain the readability of individual panels, mean observed survivals greater than 0.2
666 are not plotted.

667

668



669

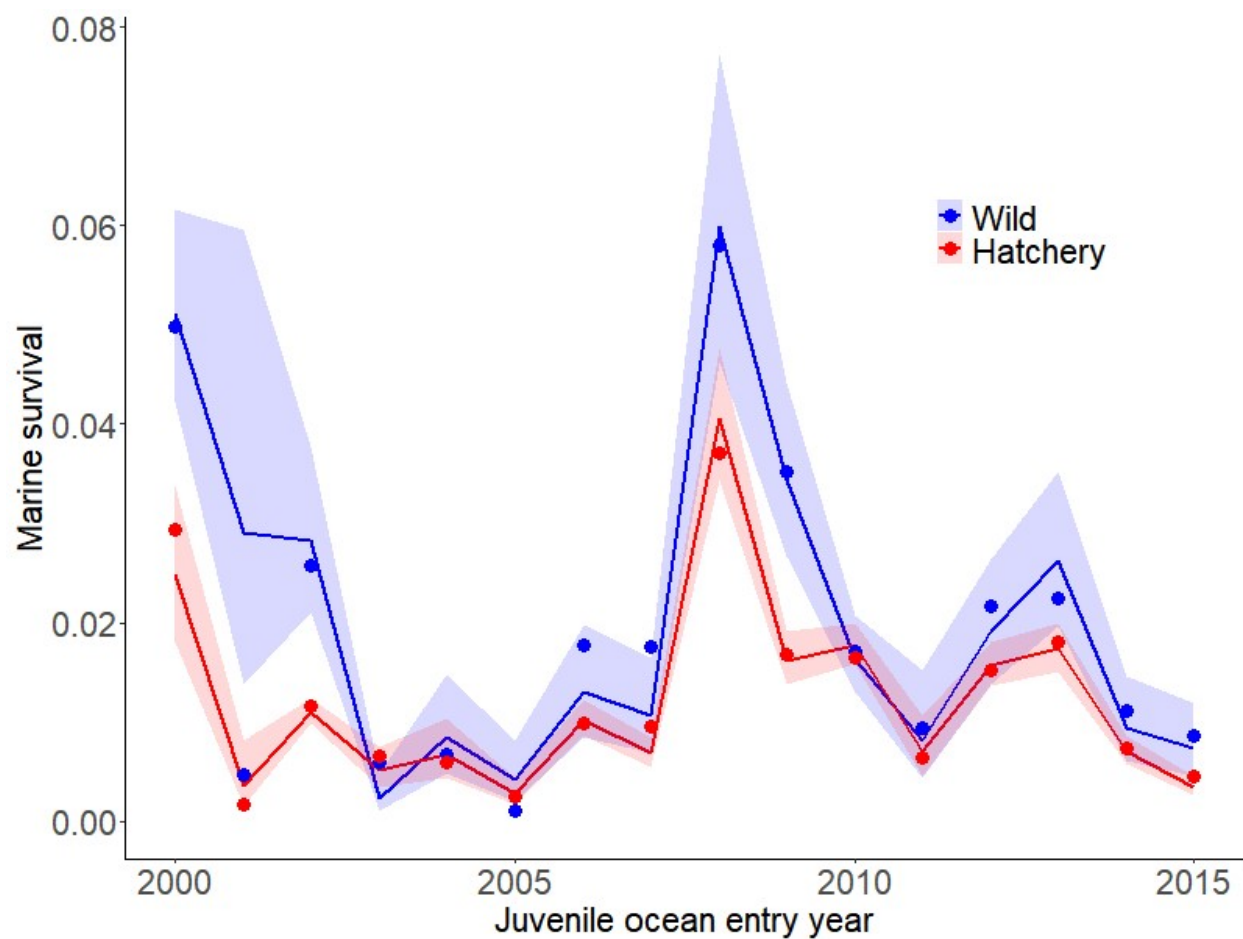
670

671 Figure 4. Predicted smolt-to-adult survival by day for hatchery (red) and wild (blue)
672 spring/summer Chinook salmon (upper panel) for best fit models for each rear type that include
673 both day and day/year interactions(Table 4). Lines represent expected survivals and shaded

674 regions represent 95% credible intervals. Average daily proportion (across all years) of smolts
675 arriving to and migrating past Bonneville Dam (2000 to 2015) (lower panel).

676

677
678



679

680 Figure 5. Observed (points) and estimated (line) annual survival with 95% credible intervals

681 (polygons) for hatchery (red) and wild (blue) spring/summer Chinook salmon from 2000 through

682 2015.

683

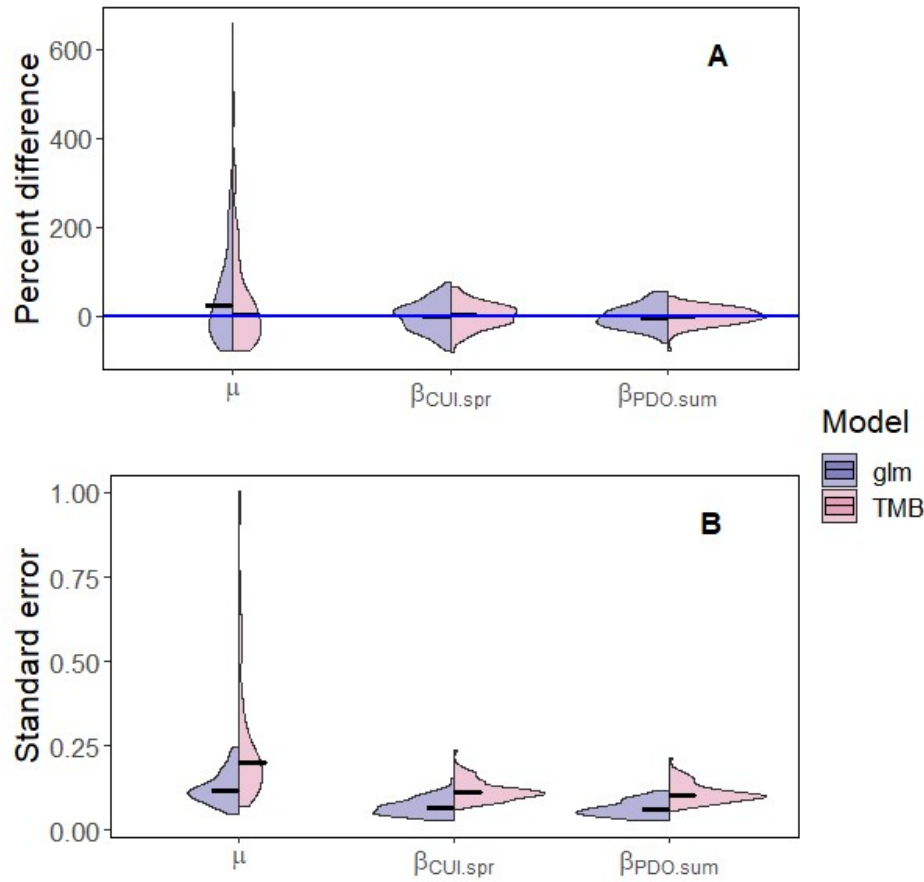


Figure 6. Split violin plot comparing the percent difference (upper panel) between the estimated and true parameter values and the standard errors for the fixed effects μ , $\beta_{CUI.spr}$, and $\beta_{PDO.sum}$ (lower panel) for the mixed-effect model with AR1 processes for day and day/year interaction (TMB; blue violins) and the fixed-effect model for day, day², and day/year interaction (glm; red violins) fit to simulated data for wild spring/summer Chinook salmon. Simulated the data were generated from the model with the lowest AIC for wild fish (see Table 4). Horizontal lines represent median values for the violins and the horizontal blue line in panel (A) represents zero percent difference between the estimated and true parameter values.

696 **Tables**

697

698 Table 1. Sample size of hatchery and wild juvenile Chinook salmon from different watersheds
699 within the Snake River Basin.

Watershed	Hatchery	Wild
Clearwater	88,895	4,512
Grande Rhonde and Wallowa	16,956	5,636
Imnaha, Pahsimeroi, South Fork Salmon	60,215	14,622
Little Salmon	69,016	8
Lower Snake and Tucannon	11,125	988
Middle Fork Salmon	0	3,714
Upper Salmon River (above Yankee Fork) and Lemhi	5,589	4,384
Total	251,796	33,864

700

701

702 Table 2. Description of the environmental variable names, the years of available data, and the
703 website location of the data.

Variable	Description	Years Available	URL / Source
CRflow.spr CRflow.sum	Seasonal Columbia River flow as measured at The Dalles Dam (USGS site 14105700)	1978-present	http://waterservices.usgs.gov/rest/DV-Service.html
CRtemp.spr CRtemp.sum	Seasonal Columbia River temperatures at The Dalles Dam (USGS site 14105700)	1997-present	http://waterservices.usgs.gov/rest/DV-Service.html
cui.win cui.spr cui.sum cui.aut	Seasonal coastal upwelling index	1946-present	http://www.pfeg.noaa.gov/products/PFELData/upwell/monthly/upanoms.mon
mei.win mei.spr mei.sum mei.aut	Seasonal Multivariate ENSO Index	1950-present	https://www.esrl.noaa.gov/psd/enso/mei/
npgo.win npgo.spr npgo.sum npgo.aut	Seasonal North Pacific Gyre Oscillation	1950-present	http://www.o3d.org/npgo/npgo.php
npi.win npi.spr npi.sum npi.aut	Seasonal North Pacific Index (index of Aleutian Low Pressure)	1899-present	https://climatedataguide.ucar.edu/sites/default/files/npindex_monthly.txt
oni.win oni.spr oni.sum oni.aut	Seasonal Oceanic Niño Index	1950-present	http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml
ersstWCoast.win ersstWCoast.spr ersstWCoast.sum ersstWCoast.aut	Seasonal sea surface temperature for coastal Washington	1900-present	https://www1.ncdc.noaa.gov/pub/data/cmb/ersst/v5/netcdf/

ersstArc.win ersstArc.spr ersstArc.sum ersstArc.aut	Seasonal sea surface temperature from Johnstone and Mantua (2014)	1900-present	https://www1.ncdc.noaa.gov/pub/data/cmb/ersst/v5/netcdf/
transport.win transport.spr transport.sum transport.aut	Seasonal measure of Sverdrup index most correlated with the temperatures in the upper 20 meters.		

704

705

706 Table 3. List of data types, subscripts, parameters (i.e., fixed effects), and random effects used to
707 model the smolt to adult survival of spring/summer Chinook salmon originating from the Snake
708 River Basin.

Type	Symbol	Description
Data	n_{rjt}	Cohort of juvenile fish of rear type r migrating past Bonneville Dam on day j in year t
	k_{rjt}	Number of fish from the juvenile cohort of rear type r migrating past Bonneville Dam on day j in year t that survived to adulthood
	x_t	Vector of environmental covariates for year t
Index	r	rearing type
	j	calendar day
	t	year
Fixed-effects	μ_s	mean survival
	β	vector of marine covariate parameters
	$\rho^{(rj)}$	correlation coefficient for day effect
	$\rho^{(rt)}$	correlation coefficient for yearly effect
	$\tau^{(rj)}$	correlation coefficient for the day effect in the day/year interaction
	$\tau^{(rt)}$	correlation coefficient for the year effect in the day/year interaction
	$(\psi^2)^{(rj)}$	variance of the day effect for rear type r
	$(\psi^2)^{(rt)}$	variance of the year effect for rear type r
	$(\phi^2)^{(r)}$	variance of the day year interaction for rear type r

Random effects	ν_{rj}	day effect for day j for rear type r
	ω_{rt}	year effect for year y for rear type r
	ε_{rjt}	day/year effect for day j, year t, and rear type r

709

710

711

Table 4. Top models for each origin type based on the random effects and number of marine covariates. The gradient is a measure of the likelihood surface for the maximum likelihood estimate, PD Hessian stands for positive definite Hessian, and N.E. stands for not estimated.

Rear type	ΔAIC	Random effect			Covariate		Gradient	PD Hessian
		Day (v_{rj})	Year (ω_{rt})	Day/Yea $r(\epsilon_{rjt})$	1st	2nd		
Hatchery	0.0000	N.E.	N.E.	X	transport.sum	npgo.sum	0.00035	TRUE
	1.0002	N.E.	N.E.	X	transport.sum	npgo.aut	0.00175	TRUE
	1.2726	N.E.	N.E.	X	transport.sum	npgo.spr	0.00116	TRUE
	2.9505	N.E.	N.E.	X	ersstArc.win	transport.sum	0.00000	TRUE
	3.0477	N.E.	N.E.	X	ersstArc.spr	transport.sum	0.00070	TRUE
	3.9997	N.E.	X	X	transport.sum	npgo.sum	0.00027	TRUE
	3.9998	X	N.E.	X	transport.sum	npgo.sum	0.00043	TRUE
Wild	0.000							
	0	X	N.E.	X	cui.spr	pdo.sum	0.00019	TRUE
	1.789							
	4	X	N.E.	X	ersstArc.spr	cui.spr	0.00000	TRUE
	1.859							
	0	X	N.E.	X	npi.sum	pdo.spr	0.00029	TRUE
	2.008							
	0	X	N.E.	X	ersstArc.win	ersstWacoast.sum	0.00007	TRUE
	2.068							
	3	X	N.E.	X	ersstWacoast.sum	pdo.win	0.00007	TRUE
	2.112							
	3	X	N.E.	X	ersstArc.win	cui.spr	0.00002	TRUE
	2.495							
	5	X	N.E.	X	cui.spr	pdo.spr	0.00009	TRUE
	3.284							
	6	X	N.E.	X	ersstWacoast.sum	pdo.spr	0.00010	TRUE
	3.306							
	4	X	N.E.	X	ersstWacoast.sum	oni.spr	0.00028	TRUE
	3.529							
	6	X	N.E.	X	ersstArc.win	npi.spr	0.00018	TRUE
	3.588							
	2	X	N.E.	X	ersstWacoast.spr	npi.sum	0.00030	TRUE
	3.949							
	2	X	N.E.	X	cui.spr	pdo.win	0.00001	TRUE

Table 5. Estimated fixed effects for the model with lowest marginal AIC values for wild spring/summer Chinook salmon, and the models with the lowest marginal AIC for hatchery fish with and without a random day effect included (N.E. means not estimated).

Parameter description	Symbol	Hatchery		Wild
		without day effect	with day effect	
Mean annual survival	μ_s	0.008 (0.006, 0.01)	0.008 (0.006, 0.011)	0.009 (0.002, 0.035)
Correlation of day effect	$\rho^{(rj)}$	N.E.	0.05 (0.003, 0.517)	0.986 (-0.992, 0.994)
Correlation of day in day/year effect	$\tau^{(rj)}$	0.955 (0.853, 0.972)	0.958 (0.835, 0.974)	0.932 (0.241, 0.963)
Correlation of year in day/year effect	$\tau^{(rt)}$	-0.067 (-0.355, 0.248)	-0.109 (-0.397, 0.229)	-0.489 (-0.707, 0.058)
Process error for day effect	$(\psi^2)^{(rj)}$	N.E.	0.134 (0.02, 0.905)	0.793 (0.243, 2.588)
Process error for day/year effect	$(\phi^2)^{(r)}$	0.58 (0.449, 0.749)	0.576 (0.442, 0.75)	0.611 (0.451, 0.826)
Effect of first marine covariate	β	0.488 (0.228, 0.747) ¹	0.478 (0.218, 0.737) ¹	0.458 (0.22, 0.695) ²
Effect of second marine covariate	β	0.547 (0.283, 0.812) ¹	0.56 (0.295, 0.825) ¹	-0.608 (-0.82, -0.396) ²

¹For hatchery spring/summer Chinook salmon, the first and second marine covariates in the model with the lowest AIC were transport.sum and npgo.sum, respectively.

²For wild spring/summer Chinook salmon, the first and second marine covariates in the model with the lowest AIC were cui.spr and pdo.sum, respectively.

727 Table 6. Deviance ratios $\left(\text{i. e.}, 1 - \frac{D_m}{D_0} \right)$ for different fixed and mixed-effects models for hatchery
728 and wild fish, where D_m is -2 times the log-likelihood of model m and D_0 is -2 times the log-
729 likelihood of the null model. The deviance ratio explains how close the model is to the best
730 model that fits the data exactly.

Model	Likelihood		Deviance ratio	
	Hatchery	Wild	Hatchery	Wild
Null	1760.4	998.3	0	0
Marine	1475.2	921.6	0.162	0.077
Day	1718.6	952.1	0.024	0.047
Day/year	1151.4	802.4	0.346	0.197
Day + day/year	1151.4	800.4	0.346	0.198
Marine covariates + day + day/year	1144.1	790.8	0.350	0.208

731

732

