COEQWAL · KEY OUTCOME METHOD DOCUMENTATION

# Sacramento River Winter-run Chinook salmon

*Reproduced as written by the Winter-run salmon team, in the team's own terminology.*

**Key outcome name:** Sacramento River Winter-run Chinook salmon

***Authors:*** *Ann-Marie Osterback[1,2*], Noble Hendrix[3], Cyril Michel[1,2], Doug Jackson[3], David Boughton[2]*

****Corresponding author:*** *Ann-Marie Osterback*

*[1] University of California, Santa Cruz, Institute of Marine Sciences' Fisheries Collaborative Program, 1156 High Street, Santa Cruz, California 95064, U.S.A.*

*[2] Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 110 McAllister Way, Santa Cruz, California, 95060, U.S.A.*

*[3] QEDA Consulting, Seattle, WA 98103, U.S.A.*

# Background

Sacramento winter-run Chinook salmon ("winter-run" *Oncorhynchus tshawytscha*) are a specific run of Chinook salmon that is endemic to the Sacramento River in California's Central Valley. As with other salmon runs in California, the winter-run population has declined over the last couple of centuries for a variety of reasons, including loss of habitat, over-harvest, and changes to the water infrastructure in California (Yoshiyama 1998). Their population decline hit a critical juncture in the late 1980s and early 1990s, leading to California listing them as endangered in 1989, the Federal government listing them as threatened in the same year, and then later uplisting them to endangered in 1994 (NMFS 1994, Meyers 2021). Since then, wildlife agencies and conservation organizations have made substantial efforts to support their population recovery, yet winter-run remain in danger of extinction (Meyers 2021).

A primary barrier to the recovery of winter-run is its unique adaptation to spring-fed rivers, combined with the development of water infrastructure in the Central Valley that blocks migratory access to those rivers and requires them to use habitat elsewhere. Juvenile winter-run migrate to the Pacific Ocean in the winter where they mature for 1-3 years before they return to spawn during winter months (Figure 1). To access their spawning grounds, they migrate hundreds of miles upstream in the Sacramento River, and then hold in the freshwater for several months before spawning in the summer (Figure 1, Vogel and Marine 1991). Winter-run are the only run of Chinook salmon in the world that spawns at the beginning of the warm season, which presents unique challenges due to their reliance on cold water for egg development and survival. Historically, winter-run could succeed at this because they would migrate far upstream to spring-fed tributaries coming off Mount Shasta and Mount Lassen, most notably the Little Sacramento River, McCloud River, Pit River, Fall River, and Battle Creek. Porous volcanic rock in this region creates numerous springs discharging cool water to these rivers year round, which supported egg survival and development of winter-run throughout the heat of summer.

However, the completion of Shasta Dam on the Sacramento River in 1945 and development of hydroelectric projects in Battle Creek blocked access to all these spring-fed rivers, thereby restricting all spawning and egg development to a new location, a single reach of the Sacramento River just downstream of Shasta Dam. Because winter-run egg survival and development is dependent on cool water in the summer months (Martin et al. 2017), operators of Shasta Dam release cool water from its stratified cold water pool, which is then re-regulated by Keswick Dam (the afterbay just downstream), to artificially maintain suitable water temperatures for spawning and egg incubation. At present, the continued existence of the last wild population of this unique endemic ecotype of Chinook Salmon is completely dependent on the success of these operations.

Unfortunately, in some years the operators cannot always maintain these cool water releases long enough or cool enough to support winter-run throughout its entire spawning and egg development season. For example, in the drought year of 2014 dam operators lost temperature control just as spawning completed, and in this and the following year an estimated 95% of eggs in the river were killed by elevated water temperature (Meyer 2021). Climate change is likely to exacerbate this problem, as more frequent and severe droughts, combined with human demand for Shasta's water, deplete the cold water pool in the reservoir, while ambient warming patterns in the river downstream continue to rise.

Water and climate also pose challenges for winter-run past the spawning and egg stages. After fry emerge from eggs in the summer and fall, they rear in the freshwater environment for several months before migrating to the Pacific Ocean the following winter to begin the marine phase of their life cycle (Fig 1). Historically, juvenile salmon had access to a diversity of off-channel and floodplain habitat that naturally inundated and supported their survival, growth, and preparation for ocean entry; but access to much of that habitat has been lost through the construction of levees, canals, and water-diversions to convey water across the state of California and protect land from flooding.

Finally, juvenile winter-run Chinook also face challenges during the smolt outmigration life stage in late winter. Higher and more variable seasonal flows in winter trigger outmigration to the Pacific Ocean and support subsequent survival of migrants on their way to the ocean (Burford et al 2025, Michel et al 2021). However, these seasonal "pulse flow" events have been dampened as water is stored behind dams in the winter for use in summer months. As a result, the availability of inundated rearing habitat for fry and pulse flows for migrants have been reduced substantially compared to historical conditions.

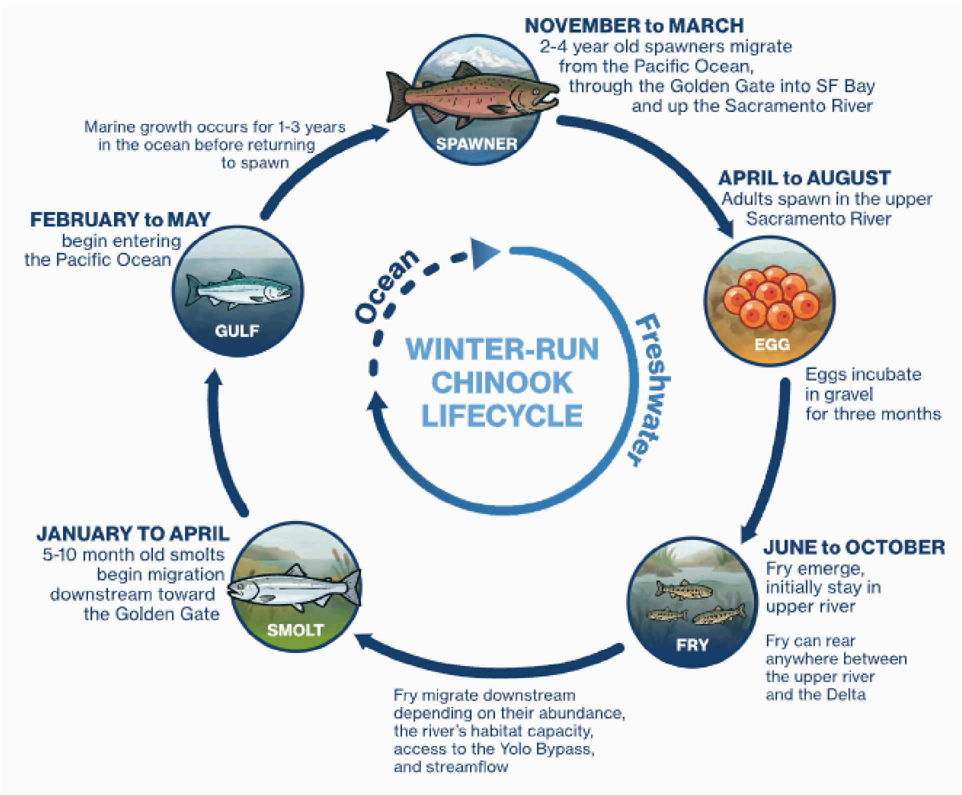


***Figure 1.*** *Life history of Sacramento River winter-run Chinook salmon.*

Because of their endangered status and reliance on re-regulated releases of cool water from Shasta Dam for egg survival, winter-run sit at the center of water operations decisions for the Sacramento River. Current operations have taken many steps to benefit winter-run, but are clearly still inadequate for the long run; and indeed have required continued operation of a fish hatchery to hedge against extinction in the wild. COEQWAL allows us to explore hypothetical water allocation scenarios that would better support winter-run recovery by supporting cool water for egg survival, increased flow to inundate rearing habitat, and larger and more frequent pulse flows to aid in outmigration survival. Such actions would surely benefit winter-run, but some key questions are: by how much? Would it be enough to substantially improve their population? The COEQWAL project provides an opportunity to explore such questions by using a quantitative life cycle model. Further, the broader COEQWAL project evaluates whether scenarios that better support winter-run are at odds with, or alternatively, may provide synergies with, other uses and needs for water throughout California.

For COEQWAL, the key outcome level we use for winter-run is based on the percentage of current spawning capacity that is filled by natural-origin female spawners (see the "Key Outcome Metric" for more information below). Higher key outcome values indicate larger population sizes that are able to better fill and utilize existing spawning habitat, whereas lower key outcome values indicate smaller population sizes.

# Locations of Interest

For the winter-run key outcomes, there is only one key outcome (see Key Outcome Metric below) and only at one location – the winter-run population in the Sacramento River. Note that for scenarios that include reintroduction of spawners to the McCloud River, spawners ultimately headed for the McCloud were still counted in the Sacramento River prior to movement upstream.

# Methodology

## The Winter Run Salmon Life Cycle Model

The primary tool for turning water allocation scenarios into recovery scenarios for salmon is the quantitative life-cycle model, with submodels for each particular life stage of interest, including eggs, fry, and outmigrating juveniles (i.e., smolts). A life-cycle model is fundamentally an accounting process, which adds up the “sum total” of survival benefits at each life-stage over the entire life-history from egg to spawning adult, and then propagates them forward to future generations. Submodels specify the way in which particular water-management actions translate to survival benefits for specific life stages, based on the best available science, and define the specific survival or movement response to management decisions.

We used the winter-run Chinook salmon life cycle model (WRLCM), previously described by Hendrix et al. (2024) and Osterback et al (2026), to estimate the number of winter-run that return to the spawning grounds for each scenario and year of the 100-year CalSim3 time series. The WRLCM is a spatially explicit state-space population model that estimates the number of individuals for each geographic area, lifestage and month in the freshwater environment, and then tracks abundances annually once fish enter the ocean phase of their lifecycle. The WRLCM integrates several submodels, specifically an egg survival model, several habitat capacity models, a river outmigration survival model, and a through-delta outmigration survival model (Sridharan et al 2022). Cumulative impacts to movement and survival are integrated over the lifecycle and multiple generations to quantify the effect on overall population dynamics. At its core, the WRLCM and submodels are all reliant on flow inputs to the model, which for COEQWAL, are generated by CalSim3. The WRLCM and its submodels have been useful for evaluating long term operations of the State Water Project (owned by California Department of Water Resources) and the Central Valley Project (owned by the US Bureau of Reclamation) in addition to two large water infrastructure projects (the Delta Conveyance Project and Sites Reservoir Project).

## Submodels

The overall WRLCM framework and submodels have been documented extensively by Hendrix et al (2024) and Osterback et al (2026), and are not repeated here. However, three submodels – the river outmigration submodel, model emulators for delta smolt survival, and the reintroduction submodel – were modified specifically to evaluate the COEQWAL scenarios and are described below.

*River outmigration.*--The positive effect of increased Sacramento River flow on salmon outmigration survival is well-researched (Michel et al 2021, Henderson et al 2019), enough to explicitly represent this relationship as a submodel within the WRLCM. Originally, the WRLCM used a flow-survival relationship based on acoustic telemetry data, where survival was empirically estimated as a function of flow on a monthly timestep (Hendrix et al 2017). However, for the COEQWAL project, monthly CalSim3 flow volumes were disaggregated to a daily time step, producing hydrographs that more closely represent the seasonal variability described by the functional flows approach (Yarnell et al. 2015). The disaggregation algorithm stitched together water-year-type-appropriate daily hydrograph patterns while preserving modeled monthly flow volumes. It prioritized minimum baseflows over peak flows and constrained flows to remain within channel-capacity limits. Of greatest importance to the river outmigration submodel is the January-May period, during which the resulting hydrographs represent wet-season baseflows, the snowmelt recession, and (when sufficient water is available) wet-season peaks that simulate storm-driven runoff events. Therefore, our previous flow-survival relationship, based on monthly flow, was too coarse for analyzing the daily hydrographs available for COEQWAL scenarios. Instead, we needed a submodel that could explore how fish respond to shaped pulse flows that unfold at a finer scale. We therefore developed a daily-resolution outmigration survival model adapted from Burford et al (2025). Their model used flow and other environmental conditions on a daily timestep to estimate outmigration survival in the mainstem Sacramento River by including relationships specific to migration initiation (e.g., pulse flows that cue fish migration, informed from outmigrant trapping data at Red Bluff Diversion Dam) and outmigration survival (informed by acoustic telemetry data).

The new submodel tracks survival at the daily timestep to capture the specific effects of pulse flows, but scales it up to the monthly timestep used by the WRLCM. We modified the model of Burford et al (2025) by 1) using a different location as the primary flow input into the model, to better align with COEQWAL standards (i.e., CalSim3 inputs from Sacramento at Wilkins Slough instead of Sacramento River at Freeport), 2) using acoustic telemetry detections at the City of Sacramento to estimate survival to this location rather than Benicia Bridge (to align with WRLCM geographic boundaries), 3) scaling daily success rates to the monthly timestep of the WRLCM by weighting survival rates by the daily proportion of fish that initiated migration (which was partially determined by daily flow at Bend Bridge), and 4) adding a random effect of individual telemetry study to the outmigration survival model, to better account for variation in the data. For more detailed information on the river outmigration submodel used for COEQWAL, see Appendix S-1.

*Delta emulators.*--The positive effects of flow on survival continue as winter-run smolts make their way through the Sacramento-San Joaquin Delta. Therefore, estimating through-delta outmigration survival is a key component of evaluating the effect of COEQWAL scenarios on winter-run. The typical submodel used by the WRLCM to capture flow-survival dynamics is the enhanced particle tracking model (ePTM, Sridharan et al. 2022), which estimates survival using a spatially explicit, agent-based model of fish movement in combination with the Delta Simulation Model (DSM2) particle tracking model (Anderson and Mierzwa, 2002). Although the ePTM is a comprehensive and powerful model that tracks survival through different pathways in the Delta, it requires significant computational power and time. This prevents the fast turnaround necessary for our application to the COEQWAL project. We therefore developed emulators to expedite our workflow. The ePTM emulators are machine learning models that mimic the complex behavior of the full agent-based model, estimating ePTM outputs for a small fraction of the computational expense. Instead of relying on the computationally expensive DSM2 hydrodynamics model, the ePTM emulators use flow inputs directly from CalSim3 to estimate through-delta survival. As a result, the delta survival submodel can run efficiently, allowing for the expedited analysis of COEQWAL scenarios. The ePTM emulators have been trained on a specific map ("grid") of the Delta and therefore cannot be applied to scenarios that have a different network of channels within the Delta. Therefore, the emulators cannot be used to evaluate the Delta Conveyance Project (DCP) scenarios, which assumes a different grid of the Delta. We were therefore unable to run the DCP scenarios as a part of the COEQWAL project. Due to current modeling limitations, the ePTM emulators did not evaluate the effects of daily-shaped hydrographs (unlike the river outmigration submodel). Instead, they relied on monthly flow outputs from CalSim. For more detailed information on the ePTM emulators used for COEQWAL, see Appendix S-2.

*Incorporating Reintroduction.*-- Prior to the construction of the Shasta Dam, winter-run spawned in the spring-fed tributaries originating near Mount Shasta, including the McCloud River. Currently, physical conditions in the McCloud River remain colder than those in the mainstem Sacramento River, particularly during the summer months when winter-run spawning occurs. Because of this high-quality habitat, a winter-run reintroduction program is already underway in the McCloud River. This watershed is identified as one of two primary locations critical for supporting the long-term recovery of the species (NMFS 2014).

To model this management action, we developed a submodel within the WRCLM to implement the reintroduction of a portion of the winter-run, which currently spawn in the mainstem Sacramento River, to the McCloud River. Winter-run spawning in the McCloud River occurs in cool, spring-fed water that is present year-round. Relative to spawning in the Sacramento River, McCloud River conditions can improve egg-to-fry survival due to reduced thermal mortality. Importantly, reintroduction requires adult collection, adult survival, juvenile collection, and juvenile survival to complete the life cycle in the McCloud. These rates are defined by the quality of the reintroduction effort and costs associated with its implementation. We utilized the WRLCM with a reintroduction submodel (WRLCM-R) to evaluate the scenarios under COEQWAL. We parameterized the reintroduction submodel such that the reintroduction quality would be high enough to achieve abundances that were equal to, or greater than, baseline levels in the absence of reintroduction. While Battle Creek represents the second primary reintroduction location in the Central Valley for winter-run and also has an active program underway, it is not included in the submodel described here. For more detailed information on incorporating reintroduction into the WRLCM for COEQWAL, see Appendix S-3.

Although we ran all COEQWAL scenarios both with and without reintroduction to evaluate outcomes for winter-run Chinook salmon, not all scenarios with reintroduction are included on the COEQWAL Data Dashboard visualization tools. Instead, only a select few scenarios with reintroduction are included to allow for comparisons across multiple outcomes, while not overwhelming the user with too large of a number of scenarios to visualize. The select scenarios that are included with reintroduction are identified with a "-R" following the scenario number (e.g., s0020-R), where "-R" indicates "with reintroduction". However, results for all scenarios, both with and without reintroduction, are available in Data page.

## Key Outcome Metric

The metric for winter-run response for each scenario is the number of female natural-origin spawners per year, as estimated by the WRLCM or WRLCM-R. Spawner abundance, averaged over multiple years, is one of the primary metrics used to evaluate population status for salmon (NMFS 2025), because these individuals survived the cumulative challenges encountered throughout the life cycle. Furthermore, females are the ones that drive reproduction rates and can be limited by available spawning habitat, and thus are the focus of the metric.

Here we focus specifically on natural-origin fish—individuals that originated from eggs laid in the Sacramento River as opposed to in a nearby hatchery. The U.S. Endangered Species Act says its purpose is "to provide a means whereby the ecosystems upon which endangered species and threatened species depend may be conserved" and the metric for this purpose here is natural-origin fish. Likewise, natural-origin fish are the focus of recovery plans (NMFS 2014), and because they spend their entire juvenile lives in the river system, they respond to changes in water allocation in ways that hatchery-reared fish do not.

The WRLCM does account for low-level contributions of hatchery-origin fish that can contribute to the natural-origin population in later generations. The WRLCM assumes that the Livingston Stone National Fish Hatchery (LSNFH) broodstock program collects up to 60 natural-origin spawning females each year, spawns them, raises their progeny at the hatchery, and eventually releases them as smolts into the Sacramento River just prior to typical outmigration timing. If these smolts successfully return as spawning adults, they lay their eggs in the Sacramento River and these eggs are then considered part of the "natural-origin" population.

For each scenario, we summarized female spawners per year into one value that simultaneously accounts for abundance relative to available habitat, varying cohort strength, and model uncertainty (Figure 2). Specifically, we defined the key outcome metric as the "**maximum percentage of spawning capacity achieved by the 3-year rolling average of natural-origin female spawners, evaluated using the lower 20th percentile of WRLCM simulations**", which we unpack as follows:

First, to calculate the percent of spawning capacity achieved by female spawners for a given year, we divided the total number of female spawners by the total spawner capacity, which was the sum of spawning habitat below Keswick and in the McCloud River (i.e., total spawner capacity was 34,000 spawners, see Appendix S-3 for more information). We maintained this same habitat capacity (i.e., 34,000 females) for all scenarios, regardless of whether reintroduction actions were included, ensuring comparability between actions. This approach allows us to use the percent of spawning capacity as a metric of relative abundance among scenarios.

Second, we used a 3-year rolling average to smooth the response by generation time, which is typically about 3 years for winter-run Chinook. This averages across fluctuations of individual cohorts (year classes) within a generation, providing a clearer signal of overall population response. We constrained the key outcome metric to the 3-year rolling average between the years 1934 and 2016. We excluded years prior to 1934 because all scenarios demonstrate an initial, steep population drop driven by the severe extended drought conditions of the late 1920s and 1930s (the Dust Bowl era). By starting in 1934, we therefore focused our metrics on scenario performance following the initial severe drought conditions. Finally, we excluded years after 2016 because the cohorts that spawned during that period have not yet completed their ocean maturation cycle and returned to freshwater to spawn by the final year of the COEQWAL time series (2021).

Third, we accounted for model uncertainty by selecting the lower 20th percentile of 1,000 WRLCM simulation "time-series traces" generated for each scenario. These traces are generated from the distribution of uncertainty in WRLCM parameter estimates. We chose the 20th percentile to take a moderately risk-averse stance, in that we wanted to be at least 80% confident that a scenario was the reported value or better.

Finally, we used the **maximum percentage of spawning capacity achieved by the 3-year rolling average of natural-origin female spawners, evaluated using the lower 20th percentile of WRLCM simulations** as the single-value key outcome to represent the winter-run population response for a given scenario. This key outcome therefore represents the three years in the time series that had the highest average number of spawners across all three dominant age-3 cohorts. Because salmon populations naturally fluctuate, we did not expect a recovering salmon population to meet key outcome targets every year of a time series. For example, during an extended drought, we would expect the population to decline, and during extended wet conditions, we would expect the population to increase. However, to identify a scenario that supports winter-run, we would expect spawners of all cohorts (using the 3-year rolling average) to meet target levels for key outcomes *at least once* throughout the 100-year CalSim3 time series. Therefore, we used the maximum value of the 3-year rolling average across the 100-year CalSim3 time series as the key outcome to determine the highest level that was achieved at least once in the time series.

*Panel A Panel B*

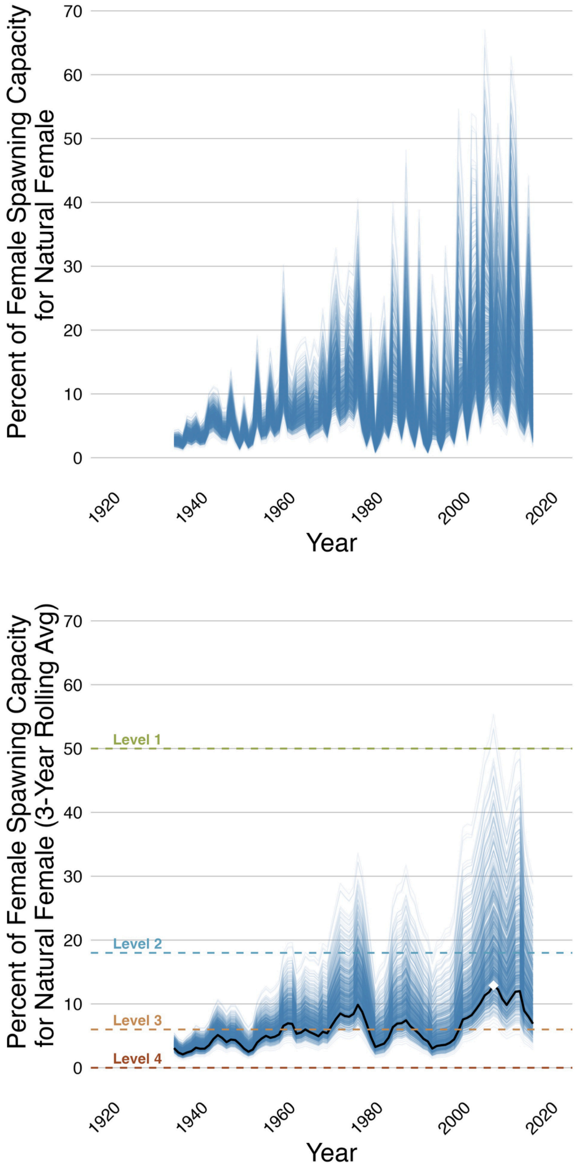


***Figure 2.*** *The progression of how the winter-run key outcome was calculated, using an example scenario. First, the percent of spawning capacity achieved by female spawners was calculated across years and the 1000 WRLCM traces generated from the WRLCM (Panel A). Second, those data were then used to generate the WRLCM key outcome metric by 1) calculating 3-year rolling average for each year (the average of the given year and the two years prior (Panel B, blue lines)), 2) extracting the lower 20th percentile simulation for each year (Panel B, black line), and 3) identifying the maximum value of that black line (white diamond), which represents the "* ***maximum percentage of spawning capacity achieved by the 3-year rolling average of natural-origin female spawners, evaluated using the lower 20th percentile of WRLCM simulations"*** *– this is the single value used to inform the key outcome metric for each scenario, and is assigned a key outcome level (Level 1, 2, 3, or 4, indicated by horizontal dashed lines, and discussed in "Discrete Key Outcome Levels" below).*

## Discrete Key Outcome Levels

To be consistent across the COEQWAL project, we converted the winter-run key outcome to a common scale to better compare with results from other COEQWAL outcomes. The key outcome was categorized into 4 different outcome levels based on the degree of spawning habitat saturation (Table 1): the winter-run population demonstrates optimal habitat saturation (Level 1), acceptable habitat saturation (Level 2), at-risk levels of habitat saturation (Level 3), or demonstrates critical levels of habitat saturation (Level 4).

***Table 1.*** *Key outcome level, description, definition and rationale.*

| Key Outcome Level | Description | Definition | Rationale |
|---|---|---|---|
| 1 | Optimal | Level 1 is met if there's at least an 80% chance (>800 out of 1,000 model runs) that the number of natural-origin winter-run females achieves **greater than or equal to 50%** of the available spawning habitat, using a rolling 3-year average. | At greater than 50% of available spawning capacity, the population occupies the majority of available habitat. This indicates high habitat saturation, where a substantial portion of the available spawning grounds is utilized. |
| 2 | Acceptable | Level 2 is met if there's at least an 80% chance (>800 out of 1,000 model runs) that the number of natural-origin winter-run females achieves **greater than or equal to 18% and less than 50%** of the available spawning habitat, using a rolling 3-year average. | At 18–50% of available spawning capacity, the population occupies a significant portion of the available habitat, though it still leaves substantial capacity available for further population expansion. |
| 3 | At-risk | Level 3 is met if there's at least an 80% chance (>800 out of 1,000 model runs) that the number of natural-origin winter-run females achieve **greater than or equal to 6% and less than 18%** of the available spawning habitat, using a rolling 3-year average. | At 6–18% of available spawning capacity, abundance is sufficient to avoid modifying commercial harvest rules, yet the population utilizes only a small fraction of the available spawning habitat. |
| 4 | Critical | Level 4 is met if the number of natural-origin winter-run females do not satisfy Levels 1, 2, or 3. This is equivalent to achieving **less than 6%** of the available spawning habitat. | Utilizing less than 6% of available spawning capacity means that habitat is critically under-utilized. This threshold corresponds to an escapement of 3,000 total spawners[1]; below this, the Winter-run Harvest Control Rule (NMFS, 2018) requires changes to commercial harvest practices to mitigate impacts on the vulnerable winter-run population. |

[1] The value of 3,000 spawners was multiplied by the observed average female sex ratio of age-3 and age-4 spawners (2003–2020) to calculate the number of females corresponding to an escapement of 3,000 individuals (Killam, personal communication), which equates to approximately 6% of the total spawning capacity.

### Continuous Key Outcome Levels

The Discrete Key Outcome Levels cannot provide insight into performance of scenarios within each of the four categories (Table 1). To address this limitation, we utilized a continuous metric to reflect the variation among scenarios within discrete levels. Therefore, we included continuous outcome level values, ranging from 1.0 - 4.99 as follows.

- 1.00–1.99 for Level 1
- 2.00–2.99 for Level 2
- 3.00–3.99 for Level 3
- 4.00–4.99 for Level 4

The decimal is related to the position of the key outcome between the upper and lower thresholds within its assigned level. Using the scenario in Figure 2 as an example, the winter-run key outcome (white diamond) is assigned to Level 3, which by definition means its percent spawning capacity is between the upper threshold (18%) and the lower threshold (6%) of Level 3 (Table 1). If the value of that key outcome is 13%, then the continuous level value would be 3.42, since it lies 42% of the distance from the upper key outcome threshold (18%), and 58% of the distance from the lower outcome threshold (6%) of Level 3. In this framework, a continuous level value of 3.0 represents the top of the Level 3 range (closest to Level 2), while 3.99 represents the bottom of the range (closest to Level 4). Using this approach, we can identify differences among key outcomes within a level, where lower continuous level values outperform higher continuous level values.

# Guidelines for Interpretation

The following considerations are critical when interpreting winter-run key outcomes and comparing performance across scenarios:

## The Shasta bottleneck and life-stage trade-offs

Winter-run are unique in their reliance on cold water habitat to prevent egg mortality during the summer. Even if scenarios support other life stages, such as providing environmental flows to support rearing and migration, those benefits will not be realized if the scenario fails to maintain the Shasta Reservoir cold water pool, or there is no reintroduction to the McCloud River, to support summer egg survival. Because other salmonids in the Sacramento River (e.g., fall-run, spring-run, steelhead) are not as dependent on summer Shasta Reservoir cold water releases, they may benefit from scenarios that winter-run do not.

Ultimately, all life stages must be supported to support winter-run recovery and thus it is important to understand how different actions affect survival at various lifestages. For example, fry productivity (the number of fry that emerged divided by the number of eggs) is influenced by water temperature during egg incubation, which can be supported by reintroduction to the cool McCloud River or by maintaining the cold water pool which can be estimated by the end of April storage in Shasta Reservoir. Similarly, smolt productivity (the number of smolts that entered the ocean divided by the number of emerged fry) is influenced by instream flows that support rearing habitat and outmigration survival and can be positively related to flow at locations such as Bend Bridge, Wilkins Slough, or Net Delta Outflow. Therefore a scenario might support one lifestage quite well, but can still result in an undesirable key outcome level because it does not support other lifestages, and therefore the cumulative survival remains low.

## Reintroduction and Holistic Recovery

Salmon recovery is not solely reliant on modifications to hydrology, but will likely also require improvements to habitat, harvest practices, and hatchery management (the 4-H's; Gore and Doerr 2000). This is supported by a recent study by Osterback et al (2026), that demonstrated a holistic approach that supported each lifestage by modifying all 4-Hs was most likely to lead to population recovery of winter-run Chinook salmon. Therefore, COEQWAL scenarios that modify hydrology alone generally do not achieve high-performing outcome levels. Therefore, additional actions such as reintroducing winter-run to historical spawning habitat in spring-fed reaches above Shasta Dam (e.g., the McCloud River), could provide consistent cool water for eggs without relying on Shasta storage. This could alleviate the pressure from the multitude of water demands from Shasta Reservoir, and allow Shasta to reallocate more water to instream flows to support winter-run outmigration and other native species. Potentially more modest changes to water allocation could further improve key outcomes if paired with habitat restoration or reintroduction. However, due to the limited amount of current spawning habitat capacity in the McCloud River, it is important to note that successful spawning will also be required downstream of Keswick, in combination with the McCloud River, to reach desirable key outcomes.

The COEQWAL results demonstrate that by adding reintroduction actions, most scenario key outcomes reach a Level 3, even under extreme future climates. In contrast, very few scenarios without reintroduction achieve outcomes better than Level 4. This highlights the importance of reintroduction in providing access to the cool water necessary for egg development and survival across a broad range of climate futures. Importantly, the reintroduction submodel that was used for these scenarios assumed reintroduction quality was quite high, meaning that the "losses" from reintroduction (e.g., inefficient capture, transport mortality and/or migration mortality or issues with routing) were minimal. Therefore, programs with lower reintroduction quality may not achieve the same outcome levels as is described here. Conversely, a more dynamic reintroduction program could potentially increase the benefit to winter-run even more than what our results describe. A flexible and dynamic program could modify the timing and number of reintroduced spawners to maximize the thermal and survival benefits that each habitat provides.

## Interpreting Level 3 and Level 4 Outcomes

As of the summer of 2026, all scenarios result in an outcome Level 3 or Level 4, indicating they still only occupy a small fraction of the available spawning habitat. However, there is significant variation within these levels, and thus the continuous level values can identify the relative performance of COEQWAL scenarios that fall within the same key outcome level. For example, the majority of scenarios still outperform baseline current operations (s0020) and thus show pathways towards improving the winter-run population. Although they still fall in lower key outcome levels, they identify water operation rules that are an improvement from current operations and better buffer winter-run from further population decline.

## Winter-run key outcomes are constrained to CalSim3 hydrology (1922–2021)

The winter-run key outcome is the result of starting the population at its contemporary population size and applying a scenario's water allocation rules over the 100-yr CalSim3 historical hydrology (1922–2021). Because the CalSim3 hydrology time series begins in 1922, the winter-run population immediately enters extended drought conditions during the 1930's "Dust Bowl". As a result, winter-run populations in the model initially drop, and much of the winter-run story here is about which COEQWAL scenarios best slow that initial population decline and help rebuild the population during the subsequent years. Therefore, we focused our key outcome metric on the time period after 1934. This provides a unique opportunity to identify operations that are more resistant to drought and can "rebound" in wet years. It is important to note, however, that future hydrology may not mirror the past, and more severe or longer droughts than are evaluated here could further impact the winter-run population.

## Hatchery Supplementation as a "Floor"

The WRLCM accounts for the LSNFH conservation broodstock program which contributes to a low-level abundance "floor" for the natural-origin population. These individuals have spent their rearing stages at the hatchery, mostly shielded from decisions associated with water management, and thus can remain at low levels even during poor freshwater habitat conditions. However, their progeny, if born in the Sacramento River, contribute to the natural-origin population. Thus, some scenarios remain at low levels of abundance without going extinct due to the hatchery population contributing low-levels of natural-origin winter-run.

# References


Anderson, J., and Mierzwa, M. (2002). DSM2 tutorial—an introduction to the Delta Simulation Model II (DSM2) for simulation of hydrodynamics and water quality of the Sacramento–San Joaquin Delta.

Burford, B. P., and Michel, C. J. (2025). Facilitated migration could bolster migrant passage through anthropogenically altered ecosystems. Ecol. Appl., 1–24. doi: 10.1002/eap.70070

Gore, M., and Doerr, P. (2000). Salmon Recovery and Fisheries Management: The Case for Dam Breaching on the Snake River. *Policy Perspect.* 7, 37. doi: 10.4079/pp.v7i2.4218

Henderson, M. J., I. S. Iglesias, C. J. Michel, A. J. Ammann, and D. D. Huff. 2019. Estimating spa- tial–temporal differences in Chinook salmon out- migration survival with habitat- and predation- related covariates. Canadian Journal of Fisheries and Aquatic Sciences 76:1549–1561

Hendrix, N., Jennings, E., Criss, A., Danner, E., Sridharan, V., & Lindley, S. T. (2017). Model Description for the Sacramento River Winter-run Chinook Salmon Life Cycle Model. *Prepared for the 2017 Biological Opinion for the California WaterFix Project in Central Valley, California.*, 1–51.

Hendrix, N., Osterback, A.-M. K., John, S., Daniels, M., Jennings, E. D., Danner, E., et al. (2024). Life Cycle Modeling Framework for Chinook salmon spawning in the Sacramento River. *NOAA Tech. Memo. NMFS* NOAA-TM-NM. doi: https://doi.org/10.25923/sj1b-xs90

Martin, B. T., Pike, A., John, S. N., Hamda, N., Roberts, J., Lindley, S. T., et al. (2017). Phenomenological vs. biophysical models of thermal stress in aquatic eggs. Ecol. Lett. 20, 50–59. doi: 10.1111/ele.12705

Meyers, E. M. 2021. Protecting a displaced species in an altered river: a case study of the endangered Sacramento River winter-run Chinook Salmon. California Fish and Wildlife Special CESA Issue: 172-188. doi: 10.51492/cfwj.cesasi.8

Michel, C. J., Notch, J. J., Cordoleani, F., Ammann, A. J., and Danner, E. M. (2021). Nonlinear survival of imperiled fish informs managed flows in a highly modified river. Ecosphere 12, 1–20. doi: 10.1002/ecs2.3498

NMFS (National Marine Fisheries Service). (1994). Endangered and threatened species; status of Sacramento River winter-run chinook salmon. Federal Register Volume 59, Number 2.

NMFS (National Marine Fisheries Service). (2014). Recovery Plan for the Evolutionarily Significant Units of Sacramento River Winter-run Chinook Salmon and Central Valley Spring-run Chinook Salmon and the Distinct Population Segment of California Central Valley Steelhead. (July), 1–427.

NMFS (National Marine Fisheries Service) (2018). Endangered Species Act Section 7(a)(2) Biological Opinion and Magnuson-Stevens Fishery Conservation and Management Act Essential Fish Habitat (EFH) Response Effects of the Pacific Coast Salmon Plan Fisheries on the Sacramento River Winter-run Chinook salmon.

NMFS (National Marine Fisheries Service). (2024). 2024 5-Year Review: Summary & Evaluation of Sacramento River Winter-Run Chinook Salmon. https://repository.library.noaa.gov/view/noaa/60883

Osterback, A.-M. K., Hendrix, A. N., Daniels, M., Danner, E., Henery, R., and Lindley, S. T. (2026). Coupling habitat, hydrological and population models to evaluate strategies to recover an endangered salmonid population. Front. Ecol. Evol., 1–21. doi: 10.3389/fevo.2026.1778333

Sridharan, V. K., Jackson, D., Hein, A. M., Perry, R. W., Pope, A. C., Hendrix, N., et al. (2023). Simulating the migration dynamics of juvenile salmonids through rivers and estuaries using a hydrodynamically driven enhanced particle tracking model. Ecol. Modell. 482, 1–27. doi: 10.1016/j.ecolmodel.2023.110393

Vogel, D. A., and Marine, K. R. (1991). Guide to Upper Sacramento River Chinook salmon life history. CH2M Hill.

Yarnell, S. M., G. E. Petts, J. C. Schmidt, A. A. Whipple, E. E. Beller, C. N. Dahm, P. Goodwin, and J. H. Viers. 2015. Functional Flows in Modified Riverscapes: Hydrographs, Habitats and Opportunities. BioScience 65:963-972.

Yoshiyama, R. M., Fisher, F. W., and Moyle, P. B. (1998). Historical abundance and decline of Chinook Salmon in the Central Valley Region of California. North Am. J. Fish. Manag. 18, 487–521.

# Appendix

## Appendix S-1. Adapting the Burford et al. (2025) initiation/passage smolt model for the SWFSC Winter-run Chinook salmon Lifecycle Model

The SWFSC winter-run Chinook salmon Lifecycle model (WRLCM) is used to estimate the long-term effects of management actions and environmental perturbations on winter-run Chinook salmon. CalSim3, a water resources planning model that provides inputs to the WRLCM, provides streamflow estimates at different nodes in the Central Valley watershed at a monthly resolution. The WRLCM is dependent on these monthly streamflow estimates, and for this reason (as well as decreased computational time), the WRLCM also runs at a monthly time step.

In recent years, there has been increasing interest in the potential application of functional flows theory to water management in California's Central Valley. Functional flows theory is pertinent to dammed and managed rivers that otherwise suffer from a homogenized and reduced hydrograph, and entails recreating the primary components of a river's annual natural hydrograph so as to restore some degree of historic conditions that the local flora and fauna evolved with (Yarnell et al. 2015). One important component of functional flows theory is shaping some large peak or pulse flows during the fall, winter, and spring period. Such pulse flows are hypothesized to help distribute rearing juvenile salmon onto food-rich floodplains and other off-channel, non-natal habitats, as well as help initiate and protect oceanward migrations of juvenile salmon. There is therefore increased interest in incorporating the effect of functional flows into the WRLCM to evaluate its impacts on long-term population growth and viability. However, flow events such as pulse flows happen on the order of a few days to a week in time, which is incongruent with the WRLCM's monthly time step.

Therefore, in order to gap this divide, SWFSC has developed a WRLCM submodel that relates migration initiation and survival to a daily-time step hydrograph through the outmigration season, to ultimately upscale to a monthly estimate of smolt survival that can be used as an input in the WRLCM. An existing migration-passage model has been recently developed for the Sacramento River by the SWFSC (Burford et al. 2025), and uses a statistical relationship between environmental conditions (including flow) and migration initiation (as measured at Red Bluff rotary screw traps), as well as a statistical relationship between environmental conditions (including flow) and migration success (as measured by acoustic-tagged salmon smolts that successfully migrate through the Sacramento River and Delta and arrive to Benicia Bridge in the San Francisco Bay). However, that model was spatially incongruent with the WRLCM for two reasons:

- Issue 1: Monthly outmigration survival in the WRLCM is applied regionally, and the two relevant regions are Red Bluff to the City of Sacramento, and Colusa to the City of Sacramento. However, the migration-passage model estimates survival from Red Bluff to Benicia. Benicia is approximately 121 river kilometers further downstream than the City of Sacramento.

- Issue 2: Functional flow simulations being generated by UC Davis collaborators provide flow at different nodes along the river, including at the Bend Bridge gaging station in the upper Sacramento River, and Wilkins Slough gaging station in the middle-lower Sacramento River. The migration-passage model uses flow inputs at both Bend Bridge (to predict migration initiation) but also uses flow at Freeport (to predict migration success) which is not a node currently being used by collaborators.

The SWFSC therefore decided to adapt the migration-passage model to overcome these spatial incongruencies.

The migration initiation submodel of the migration-passage model did not need to be modified from the original (Burford et al. 2025) version. However, two key changes were made to the migration success submodel: 1) the model was re-fit to survival data to the City of Sacramento instead of Benicia, and 2) the model used flow at Wilkins Slough gaging station as a dependent variable instead of flow at Freeport. Flow was assigned per fish by calculating the mean flow for the period of time it was present, or would have been present (in the case of mortality), in the Sacramento River (see Burford et al. 2025 for more details).

As part of the original Burford et al. (2025) migration success submodel, model selection was used to identify the most important environmental variables and most parsimonious model structure. The independent variable used in this binomial Generalized Additive Model (GAM) was the presence or absence of fish at the Benicia receiver array. As described in Burford et al. (2025), receivers do not have perfect detection and typically, survival analyses are done in a Cormack-Jolly-Seber (CJS) model framework that estimates both detection probability and survival jointly using maximum likelihood estimation. However, the CJS model framework is not well suited to nonlinear statistical methods such as GAM models. Therefore, in Burford et al. (2025), CJS models were first run for each release group of acoustic tagged juvenile salmon, and only release groups that had an estimated detection efficiency at the Benicia array equal to or upwards of 80% were included in the GAM model.

Similarly, when we adapted this model to estimate survival to the City of Sacramento receiver array, we only used release groups with near perfect detection efficiency (>=85%) as estimated by CJS models to minimize any biases caused by not formally accounting for detection efficiency in the GAM. In order to ensure estimated survival rates are representative of the winter-run outmigration corridor, we only included release groups of acoustic tagged juvenile Chinook salmon that were released at least 345 river kilometers upstream of the Golden Gate bridge in the Sacramento River (or shortly upstream in a tributary of the Sacramento River).

One noteworthy complication is that during high flow events on the Sacramento River (upwards of ~25,000 cfs), a portion of the Sacramento River spills over the Fremont Weir into the Yolo Flood Bypass, which shunts around the City of Sacramento to minimize flood risk. During these uncommon events (9.0% of fish in the acoustic telemetry dataset were migrating while Fremont Weir was spilling), any acoustic tagged juvenile salmon in the Sacramento River could go into Yolo Bypass and therefore not be detected at the acoustic receivers in the City of Sacramento despite surviving through the Sacramento River, thereby biasing survival low for that release group. To mitigate for this issue, we made two adjustments, one to the dataset and one to the GAM. Firstly, we considered any fish detected on any receiver in Yolo Bypass as a "successful" migrant, same as a fish detected at the City of Sacramento receiver array. In other words, these two receiver arrays together constitute a synthetic receiver array that spans all possible routes down the Sacramento River at this latitude, similar to the analysis in Michel et al. (2021). Second, we modified the GAM structure for the migration success submodel found in Burford et al. (2025) to include a random effect per telemetry study, so as to explain unexplained variations in migration success not explained by the fixed effects (such as survival biases at higher flow values).

The final dataset of acoustic tagged Chinook salmon included 66 distinct release groups spanning the years 2013 to 2024 (absent 2015), with detection efficiency at the City of Sacramento ranging from 85% to 100% (mean across studies 97.6%), for a total of 17,888 individual acoustic tagged juvenile Chinook salmon. These fish were released as early as November 29$^{th}$ to as late as June 21$^{st}$ within the calendar year (mean release date March 13$^{th}$), experienced flow rates at the Wilkins Slough gaging station ranging from 3342 cfs to 27057 cfs (mean 9607 cfs), and ranged in fork length from 74 mm to 215 mm (mean 98 mm).

In Burford et al. (2025), model selection was used to identify the most parsimonious migration success model that explains survival to Benicia. In this new effort, we are estimating survival to the City of Sacramento as a function of environmental variables, but we did not perform a new model selection exercise. Survival from release to Sacramento is a major component of the overarching survival from release to Benicia (the former span representing on average 73% of the river length of the latter span), and as a result, these two survival metrics are highly correlated (correlation coefficient of 0.81). We therefore presume that variables that best describe survival to Benicia will also effectively describe survival to the City of Sacramento. The GAM model structure for the migration success model from Burford et al. (2025), and adopted for this new effort, was therefore arrival for individual fish to Sacramento or Yolo Bypass (i.e., migration success) as a function of 1) fish length (mm), 2) release distance upstream (in river kilometers), 3) julian day of the water year, 4) flow at the Wilkins Slough gaging station (cfs), and 5) a random effect of study (33 unique studies in total).

Once the migration success submodel was fit to the new dataset, we created a function in R that uses the parameter estimates from both submodels to make predictions as a function of user-defined covariates. As it pertains to the migration initiation submodel (from Burford et al. 2025), we only used the "smolt" submodel for the function (a fry initiation model was also described in that publication). The smolt migration initiation submodel will therefore predict how many smolts will begin their migration as a function of user-defined flow and day of year. The migration success submodel will predict what proportion of smolts will survive their migration to Sacramento as a function of user-defined flow, fish length, day of year, and release distance. The function requires all covariates to run such that both submodels will be using the same day of year and flow data to predict both metrics. Finally, the function then takes the daily product of both predictions (i.e., number of smolts that initiate migration multiplied by survival of those smolts during migration), such that the result is the number of fish that survived to the City of Sacramento for each day. The model will make daily predictions for the length of days a continuous hydrograph is supplied. Due to the temporal bounds of the acoustic telemetry data informing the model, the function was designed to only make daily predictions from December 1$^{st}$ through June 30th. This is inconsequential to the WRLCM as those dates bound the entire outmigration window of winter-run Chinook salmon.

Finally, the function needs to upscale predictions to the monthly timestep that the WRLCM runs on. Therefore, we used a weighted mean approach where for each month, a survival prediction is generated by summing all the daily number of migrants predicted to have survived to Sacramento for that month, and dividing by the sum of all the daily number of migrants predicted to have initiated their migration for that month. One final relevant nuance of the WRLCM is that outmigration survival is needed for a few different geographic regions. For the Sacramento River, these regions are Salt Creek (i.e. Red Bluff) to the City of Sacramento, and Colusa to the City of Sacramento. The function can produce these two survival estimates by changing the release distance input from the migration success submodel. Therefore, for the Salt Creek to City of Sacramento region, the release distance input should be set to 462 kilometers, and to 307 kilometers for the Colusa to City of Sacramento region.

One final caveat worth discussing regarding the appropriateness of this model for use in the WRLCM is that the migration-passage model does not account for the amount of fish available to initiate for migration. In other words, the migration initiation submodel of the migration-passage model could (in rare instances) predict that more fish will initiate their migration on a given day than there are fish left to move, especially later in the outmigration season. However, the migration initiation submodel predictions are only used for the relative weighting of daily survival rates within a month, and therefore any overpredictions are inconsequential. Furthermore, the WRLCM does track the number of fish that initiate migrations per month independently, and would only be using the weighted monthly survival prediction from the migration-passage model.

## References

Burford, B. P., J. J. Notch, W. R. Poytress, and C. J. Michel. 2025. Facilitated migration could bolster migrant passage through anthropogenically altered ecosystems. Ecological Applications 35:e70070

Michel, C. J., J. J. Notch, F. Cordoleani, A. J. Ammann, and E. M. Danner. 2021. Nonlinear survival of imperiled fish informs managed flows in a highly modified river. Ecosphere 12:e03498.

Yarnell, S. M., G. E. Petts, J. C. Schmidt, A. A. Whipple, E. E. Beller, C. N. Dahm, P. Goodwin, and J. H. Viers. 2015. Functional Flows in Modified Riverscapes: Hydrographs, Habitats and Opportunities. BioScience 65:963-972.

# Appendix S-2. Emulator development for the delta smolt survival model

## Introduction

Estimates of smolt outmigration survival through the Delta are an essential component of the winter-run life cycle model (WRLCM). In the current version of the WRLCM, survival estimates for smolts entering the Delta via the Yolo Bypass (Yolo) and Sacramento River (River) and for smolts originating from Delta-rearing fry (Delta) are provided by the ePTM. The ePTM is a spatially explicit, agent-based model of fish movement and survival that extends the DSM2 particle tracking model (PTM) by adding swimming behaviors and a mortality model. Like the DSM2 PTM, the ePTM relies on hydrodynamic inputs generated by the DSM2 HYDRO module.

The ePTM requires substantial computational resources, and these can become prohibitive for projects that involve assessment of various hydrologic and operations scenarios across multiple decades, which is a typical application of the WRLCM.

For example, the scenarios evaluated in the current project span 100 years, and each year requires 8 monthly survival estimates (January through June, November, and December), for a total of 800 year-months. To estimate outmigration survival for Delta-rearing fry, virtual fish (vFish) must be released from all 367 DSM2 nodes upstream of Chipps Island. The total number of ePTM runs required to evaluate a single scenario is thus 800 × 367, or 293,600. Each ePTM run takes on the order of one minute, so a single scenario evaluation would require approximately 293,600 minutes, or 204 days, of CPU time. While this can be mitigated by running simulations in parallel, e.g., using cloud computing services or high-performance computing clusters, these strategies become very costly as the number of scenarios increases.

To address this runtime issue, we have developed ePTM emulators, which are machine learning models that can estimate ePTM outputs for a small fraction of the computational expense. Once trained, the emulators can generate River, Yolo, and Delta survival estimates for a full scenario in less than a minute on a typical desktop machine.

In the following, we describe the emulators, how they were trained, and their performance.

## Full-ePTM workflow

To explain the emulator's role, we will first describe the normal workflow used for the full (i.e., not emulated) ePTM. The ePTM requires outputs from a sequence of models and scripts, starting with CalSim3. CalSim3 outputs monthly flow values at various locations throughout the Delta, and these are in turn used as inputs to the DSM2 HYDRO module. However, the raw CalSim3 outputs are not in the form required by HYDRO, so they must first be preprocessed using various scripts provided with DSM2. These scripts are launched using a batch script, prepro.bat. After DSM2 HYDRO is run, ePTM uses the HYDRO tide file (HDF5) as an input. For the River and Yolo smolts, the full sequence is CalSim3 → prepro.bat → DSM2 HYDRO → ePTM → survival estimates. For the Delta-rearing fry, there is an additional step in which the survival estimates for all nodes are combined using a habitat capacity-based reweighting scheme, i.e., CalSim3 → prepro.bat → DSM2 HYDRO → ePTM → reweighting → survival estimates.

The most computationally expensive step in this workflow is running the ePTM. DSM2 HYDRO takes less time, but still involves substantial setup and management of large tide file outputs.

The emulators replace both HYDRO and the ePTM for the River and Yolo smolts, and in addition replace the reweighting step for the Delta-rearing fry. The emulator-based workflow is CalSim3 → prepro.bat → emulator$_i$ → survival estimates$_i$, where $i$ is the outmigration origin (River, Yolo, or Delta).

## Emulator implementation

There are various machine learning models that could be used to emulate a model like the ePTM, such as neural networks, boosted regression trees, and Gaussian process regressions[1], to name a few. During initial development of the emulators, we tested generalized linear model (GLM), random forest regression, and artificial neural net (ANN) emulators, but we ultimately decided to use Gaussian process (GP) emulators. The primary advantage of a GP emulator for our purposes is that it is easy to obtain an uncertainty estimate, which is central to our training procedure, as described below. During our exploratory analyses, the GP also had superior or comparable predictive performance relative to the alternatives.

The emulators were implemented within Python using the scikit-learn module. Specifically, we used the GaussianProcessRegressor class to create a GP with a Matérn kernel and a WhiteKernel for a nugget (noise) term. The GP was incorporated into a pipeline with a StandardScaler, which standardizes the inputs to zero mean and unit variance.

## Emulator training method

The emulator training procedure was designed with the expectation that the input parameter space would be unknown and expanding over time. The input parameter space is defined by the CalSim3 outputs generated for the scenarios of interest. For example, a single scenario will encompass a particular range of monthly flows in the Sacramento River at Freeport. An emulator trained on examples (training data points) from this scenario would be expected to perform well within this range of monthly flows (and associated values for other parameters), but may not perform well on another CalSim3 scenario with flows outside this range.

Since additional scenarios of interest will be generated over time, the full extent of the parameter space was not known a priori, and thus it was not possible to determine the number and ranges of training examples that would be required to adequately train the emulators. Therefore, we devised an adaptive training procedure in which examples were progressively added to the training data set until a specified uncertainty threshold was met across the entire parameter space defined by the scenarios available at the time. When additional scenarios were generated, we assessed the emulator uncertainty across its parameter space to determine if the emulator training covered the new scenario. If not, we progressively added training points to the emulator until the uncertainty threshold was satisfied again.

Through this adaptive method, our training data set evolved in a targeted manner to ensure that the emulator uncertainty was satisfactory across the relevant parameter space without having to guess at the appropriate parameter ranges or necessary size of the training data set.

We defined emulator uncertainty as the relative standard deviation, or coefficient of variation. For a given set of inputs (parameter values), a GP will return both a mean estimate and a standard deviation. The relative standard deviation is the ratio of the standard deviation to the mean, i.e., σ/μ. The relative standard deviation threshold was set to 0.01, which was found during testing to strike an acceptable balance between emulator performance and training effort.

The ePTM is a stochastic model, and therefore contributes additional uncertainty to the survival estimates. This uncertainty can be addressed using a nugget term in the GP, which accounts for random noise in the model output for a given set of parameter values, i.e., the run-to-run variability inherent to the ePTM.

However, in the context of defining the emulator uncertainty for training purposes, this additional source of uncertainty is problematic because it creates a floor on the relative standard deviation that does not necessarily decrease with additional training. For purposes of determining whether a region of parameter space is well covered by the training data – which is the fundamental goal of the relative standard deviation threshold – the nugget term is not desirable.

To calculate a well-behaved uncertainty estimate while still accounting for run-to-run variability, we employed the reinterpolation method[2]. This method involves the creation of an initial GP with a nugget term fit to the training data. Then, a second GP without a nugget term is created using the mean values at the training locations from the first GP. This results in a GP emulator with an uncertainty that tends to zero as training data are added.

Training proceeded in an iterative manner following these steps:

1. randomly select 25 examples from the parameter spaces of the available scenarios; an example is a year-month combination from a particular scenario, which will define a set of preprocessed CalSim3 monthly flows
2. run the full ePTM on the examples and obtain the survival estimates, either directly from the outputs (River and Yolo) or after reweighting (Delta)
3. fit a GP with a nugget term to all available training data
4. reinterpolate by fitting a second GP without a nugget term to the means of the first GP at the training locations
5. assess the emulator (second GP) uncertainty across all available parameter combinations (scenario, year, and month combinations) and identify examples with uncertainties greater than the relative standard deviation threshold
6. randomly select 10 examples that exceed the uncertainty threshold
7. repeat steps 2-6 until there are no available examples that exceed the uncertainty threshold, i.e., emulator uncertainty is satisfactory across the entire parameter space defined by the currently available scenarios

The three emulators (River, Yolo, and Delta) were trained sequentially, but the ePTM runs used to generate training data for the first emulator were reused to initialize the second emulator, and ePTM runs from training the first two emulators were used to initialize the third emulator. As a result, some scenario-year-month combinations in the training data are shared across all emulators, but additional combinations were simulated as necessary for each emulator to achieve the desired uncertainty level.

## Emulator training data

As described above, training of the emulators proceeded in an adaptive manner, with additional training data being incorporated as necessary when new scenarios became available. The following description of the training data, results, and performance of the emulators applies to just the initial round of training. Subsequent rounds of training increased the total number of examples used in each of the three emulators. Performance on the out-of-sample test data set described below generally stayed the same, or improved slightly, as additional training data were included and the emulators were retrained. Thirty complete scenarios were available for the initial round of training. Each scenario comprised 100 years and 12 months/year (for training, we can use all of the available months rather than only the months used by WRLCM), for a total of 12 × 100 × 30 = 36,000 available examples.

To generate survival estimates for a given month, 50 vFish are inserted uniformly across the release month and the ePTM is run for a total of four simulation months. The fates of the vFish are primarily determined by the hydrologic conditions in the first month, but transit through the Delta can take longer than a month for some vFish, in which case their fates will also be affected by hydrologic conditions in months 2-4, albeit to a lesser degree. However, for simplicity we only used the CalSim3 outputs in the first months as emulator inputs.

Hydrodynamics, and consequently vFish fates, are also affected by additional DSM2 HYDRO inputs beyond the preprocessed CalSim3 outputs. For example, HYDRO takes time series of gate operations as inputs. These inputs were also excluded in this version of the emulators.

Because there are additional sources of variation that are not explicitly included in the emulators – e.g., flow conditions beyond the first month and various time series inputs to HYDRO – there is a limit to the potential accuracy of the emulators. However, we judged that the preprocessed CalSim3 outputs in the release month would provide sufficient information to predict survival with an acceptable balance of accuracy and emulator complexity.

We used the following subset of preprocessed CalSim3 outputs as emulator inputs (DSS variable names optionally followed by a shorthand used in the Python code in parentheses): D_SJR028_WTPDWS (COSMA1), C_SAC041, NDO, TOTAL_EXP, C_CLV004 (calaveras), C_CSM005 (cosumnes), C_DMC000_TD (cvp), SR_60N_MOK019 (mok_ds_sr1), C_MOK022 (moke), C_SAC048 (sac), C_SJR070 (vernalis), and C_CSL005 (yolo). In addition, an export/inflow ratio (EI) was derived as (cvp + swp)/(sac + moke + vernalis).

## Training results and emulator performance

The Delta, Yolo, and River emulators were initially trained on 2640, 3090, and 3214 examples, respectively.

Emulator validation was performed by comparing the emulator output to both the training data set and an out-of-sample data set.

When fitting and evaluating a model, the out-of-sample data set is typically much smaller than the training data set. For example, a typical approach is to set aside 30% of the available observations to use as test data that are not incorporated in the training process at all.

In contrast, we had a large amount of available test data relative to the training data, for a few reasons. First, the number of training examples was driven by the uncertainty target and the adaptive training process, so the training data were generated on-demand as opposed to being a subset of an existing data set. Second, GP storage requirements and evaluation time increase rapidly with the number of training examples, which sets a practical limit on the number of training examples that can be included. Third, ePTM outputs for all scenario-year-month combinations were independently generated outside of the training workflow, providing a library of out-of-sample observations that was much larger than the number of training examples required to meet the uncertainty target and the number of examples that can be practicably included in a GP training data set.

Since a full set of ePTM outputs had been generated independently of the training workflow, we had access to multiple data sets for comparison and validation. For the scenario-year-month combinations in the training data set, we had: ePTM survival estimates generated during training; ePTM survival estimates generated independent of the training workflow; and survival estimates from the emulator. For the out-of-sample scenario-year-month combinations, we had both emulator estimates and ePTM estimates generated independent of the training workflow.

With these data sets in hand, we were able to measure emulator performance on both the training data set and a large out-of-sample data set. In addition, we were able to compare independent realizations of the ePTM at the training points, allowing us to compare the accuracy of the emulator to the run-to-run variability inherent to the ePTM.

As shown in Figure 1 and Figure 2, the accuracy of the River and Yolo emulators on the training data set was nearly equal to the variability that would be expected from repeated runs of the ePTM. In other words, the ePTM survival estimates predicted by the emulators are nearly as accurate as possible given the inherent variability of the stochastic ePTM output. The picture is somewhat different for the Delta emulator (Figure 3) because of the reweighting. The Delta survival estimate for a given scenario-year-month is a weighted average of all insertion locations upstream of Chipps Island. This reweighting process averages out the stochasticity of the ePTM, so repeated realizations of the Delta survival estimates from the full ePTM are much less variable than replicate realizations of the River and Yolo survivals. Therefore, the run-to-run variability results in a root mean square error (RMSE) of 0.017 versus the emulator RMSE of 0.031. The Delta emulator exhibits acceptable accuracy, but it is not quite as accurate at predicting ePTM outputs as a replicate run of the ePTM would be.

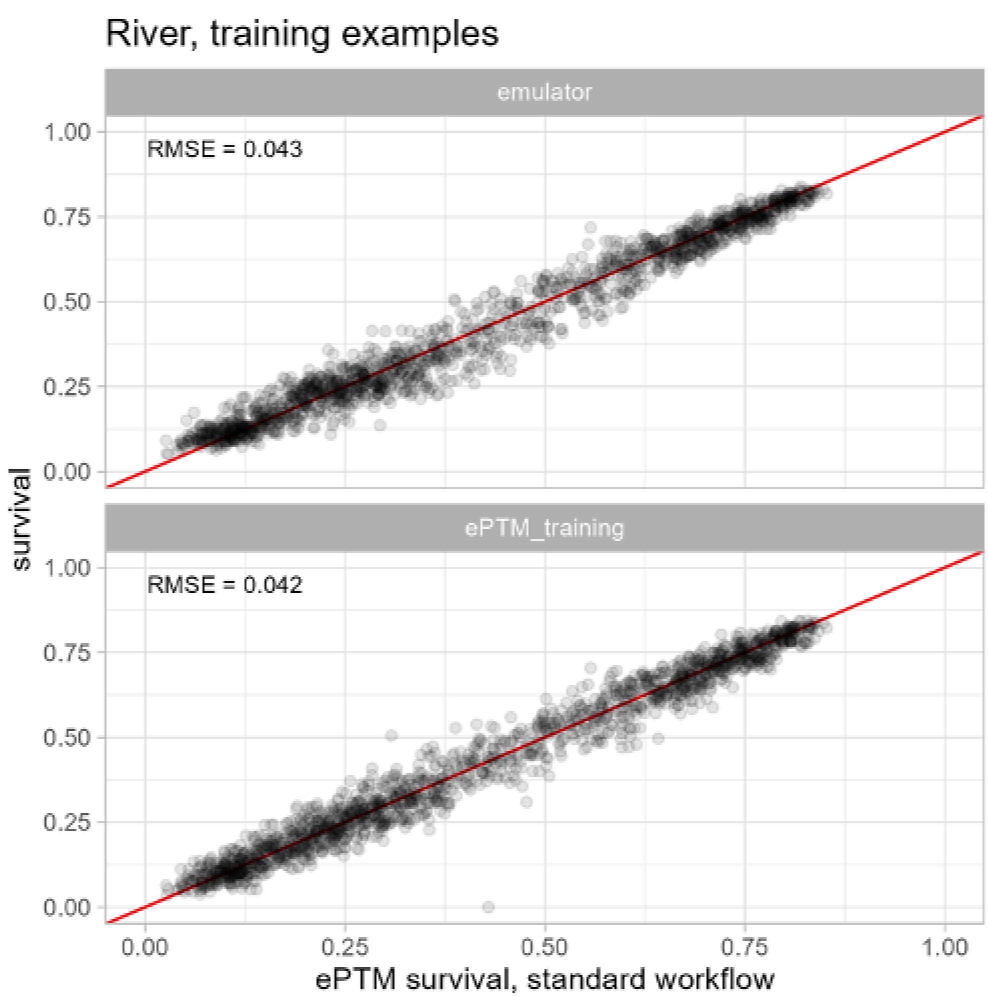

***Figure 1.*** *Performance of River emulator on training data (top); inherent run-to-run variability in ePTM outputs at the training points, as shown by comparing independent realizations of the ePTM (bottom)*

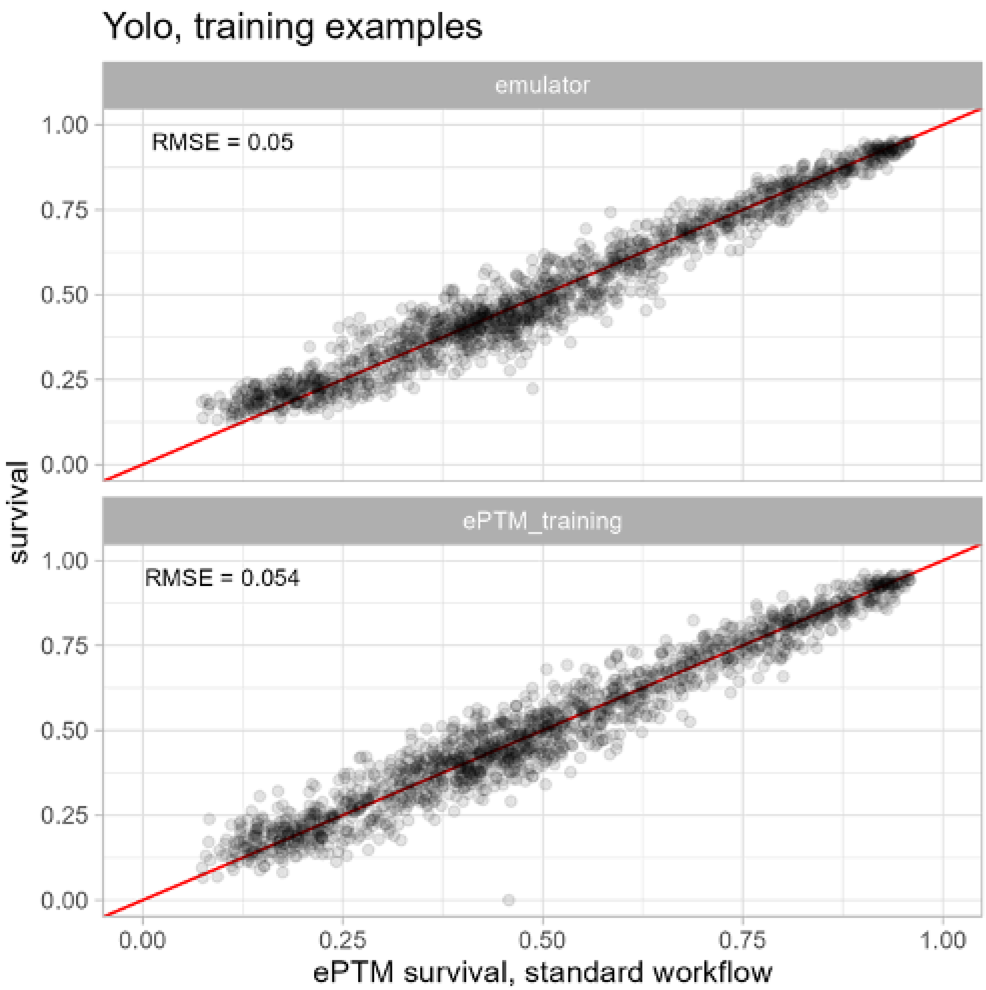


***Figure 2.*** *Performance of Yolo emulator on training data (top); inherent run-to-run variability in ePTM outputs at the training points, as shown by comparing independent realizations of the ePTM (bottom)*

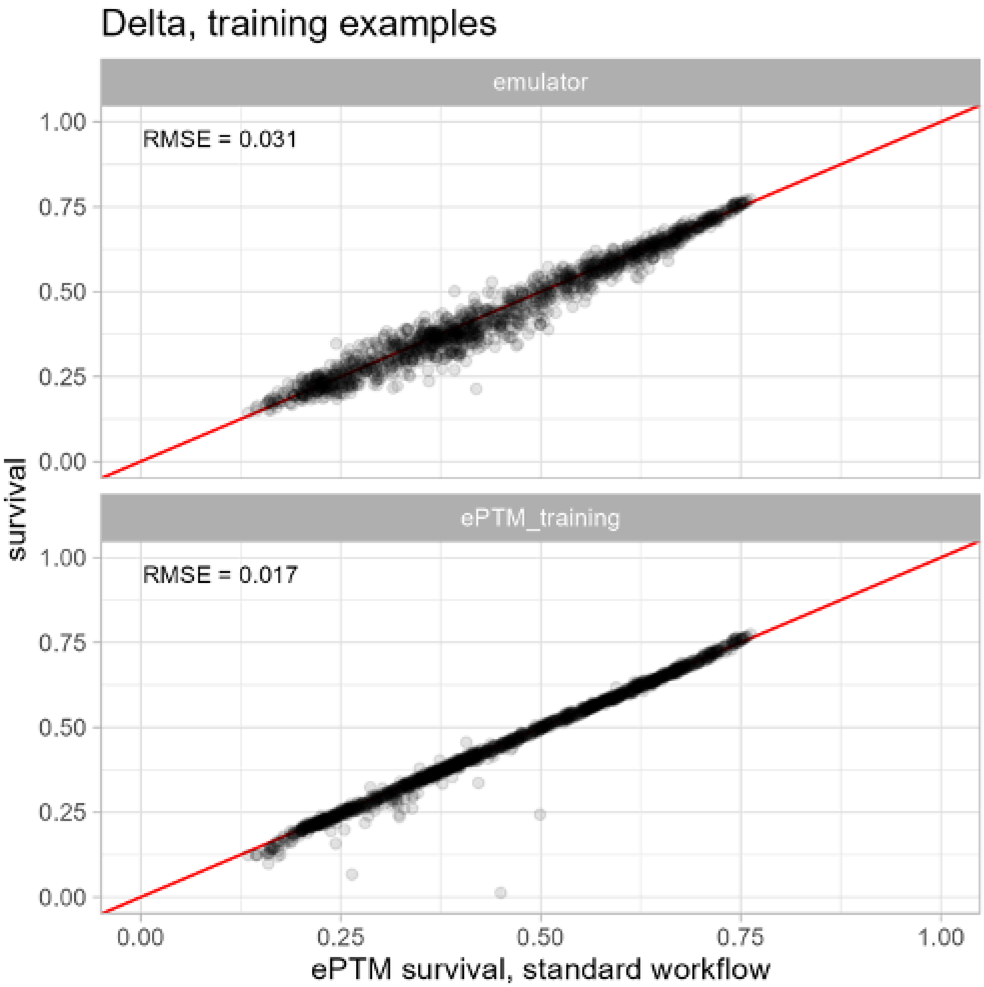


**Figure 3.** Performance of Delta emulator on training data (top); inherent run-to-run variability in ePTM outputs at the training points, as shown by comparing independent realizations of the ePTM (bottom)

The emulators' predictive performance on out-of-sample observations was generally quite similar to the performance on the training data, suggesting that the GPs are not overfitting the data (Figure 4 through Figure 6). For the River emulator, the RMSE on the out-of-sample data was 0.048 versus 0.043 for the training data; for the Yolo emulator, the RMSEs were 0.054 versus 0.05; and for the Delta emulator they were 0.035 versus 0.031.

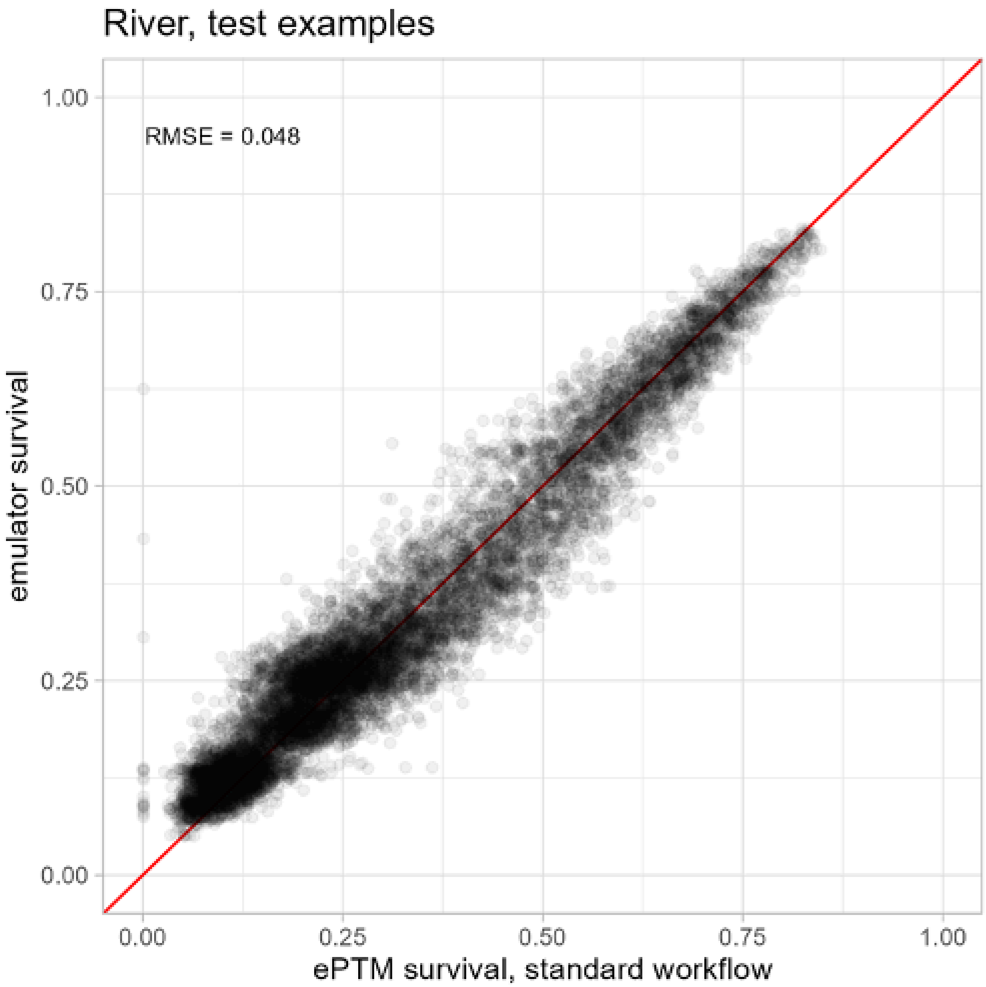


***Figure 4.*** *Out-of-sample performance of the River emulator*

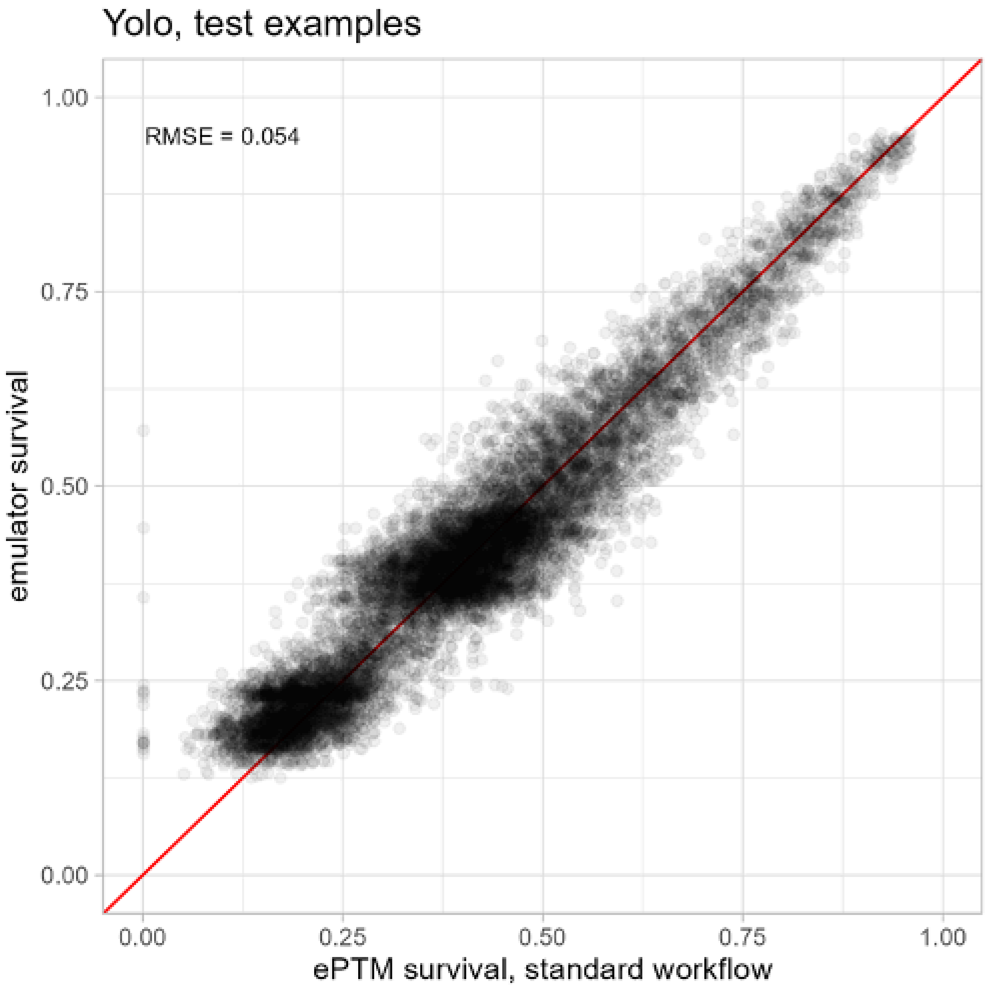


***Figure 5.*** *Out-of-sample performance of the Yolo emulator*

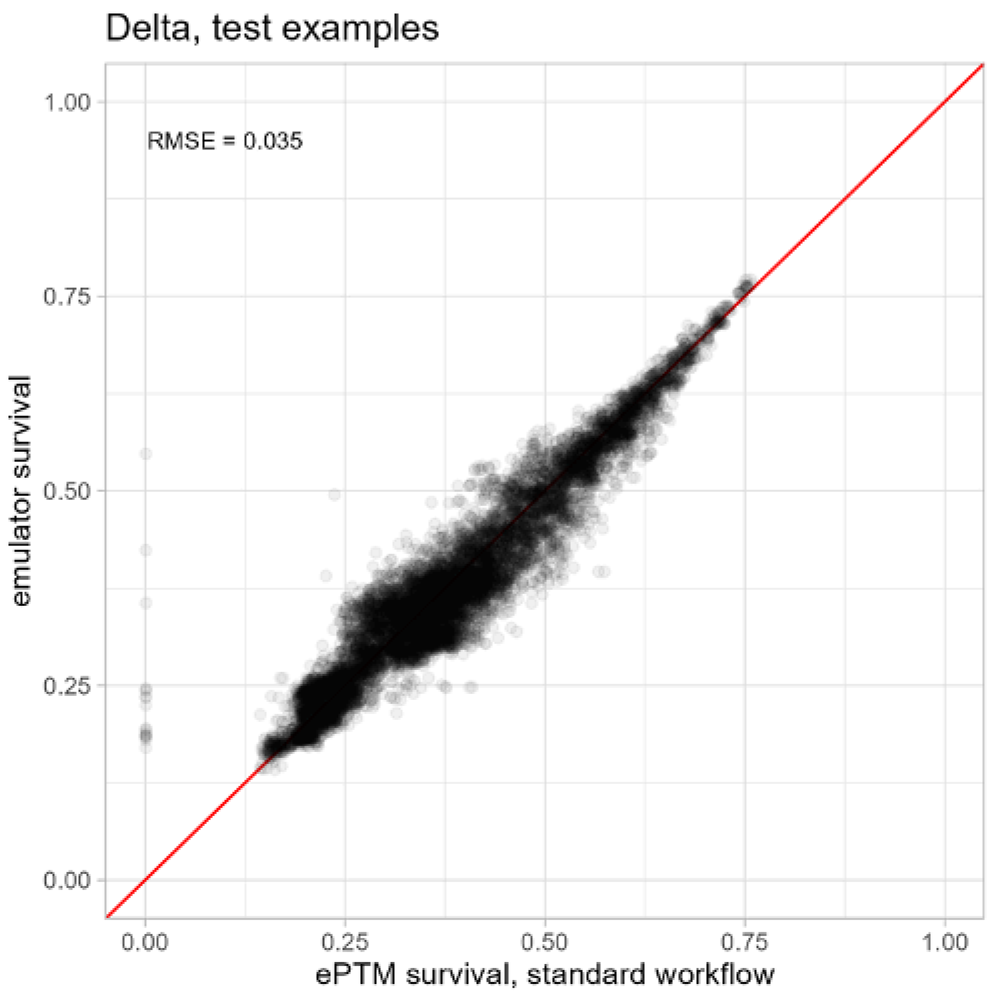


***Figure 6.*** *Out-of-sample performance of the Delta emulator*

Looking at the histograms of the residuals, there is no indication of a systematic bias: the distributions of residuals for all three emulators are centered around zero (mean = 0) and closely follow a normal distribution (Figure 7).

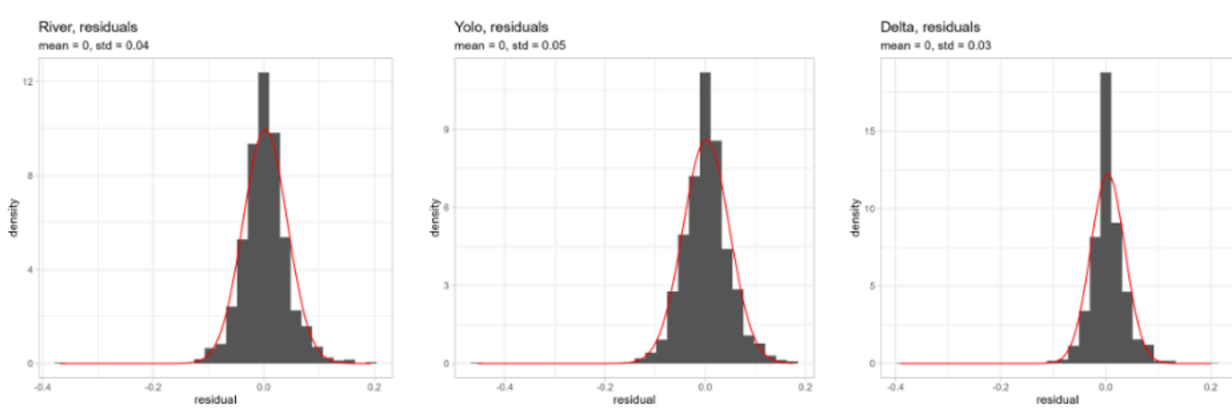


***Figure 7.*** *Histograms of residuals for River, Yolo, and Delta emulators with normal distributions overlaid; means and standard deviations of the residuals are shown in the plot titles*

There is a tendency for the emulators to overestimate the lowest survivals (Figure 8). This is a consequence of the emulators needing to fit both high and low survivals for the same input values. To illustrate this phenomenon, we created a simplified, one-variable emulator that predicts survival as a function of Sacramento River flow. As can be seen in Figure 9, survival ranges between approximately 0 and 0.25 at the lowest flow values, so the emulator splits the difference with a mean output of approximately 0.12. As a result, the emulator overestimates survival for the lowest survival values; this appears as a high bias at low survivals. It is possible that this effect could be reduced by including additional predictor variables, such as CalSim3 outputs in months 2-3.

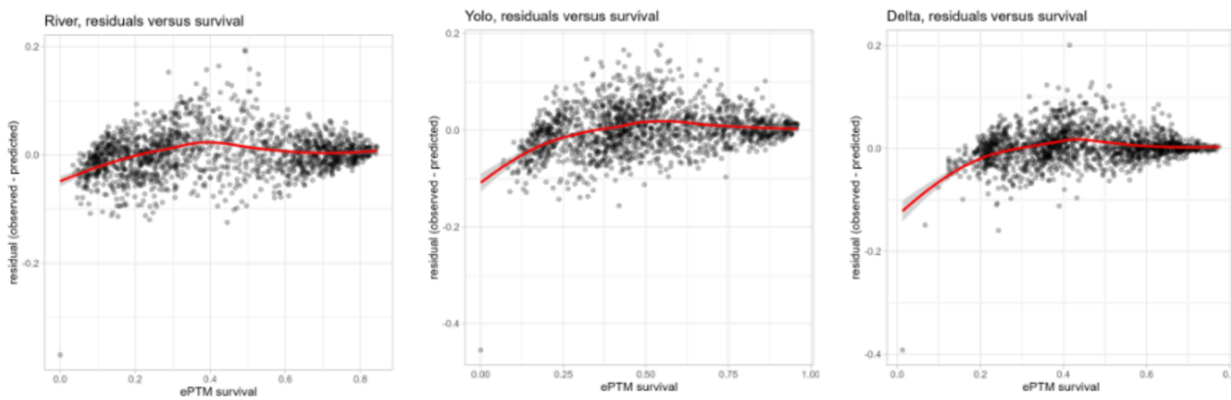


***Figure 8.*** *Residuals versus survival probabilities; red lines are LOESS curves*

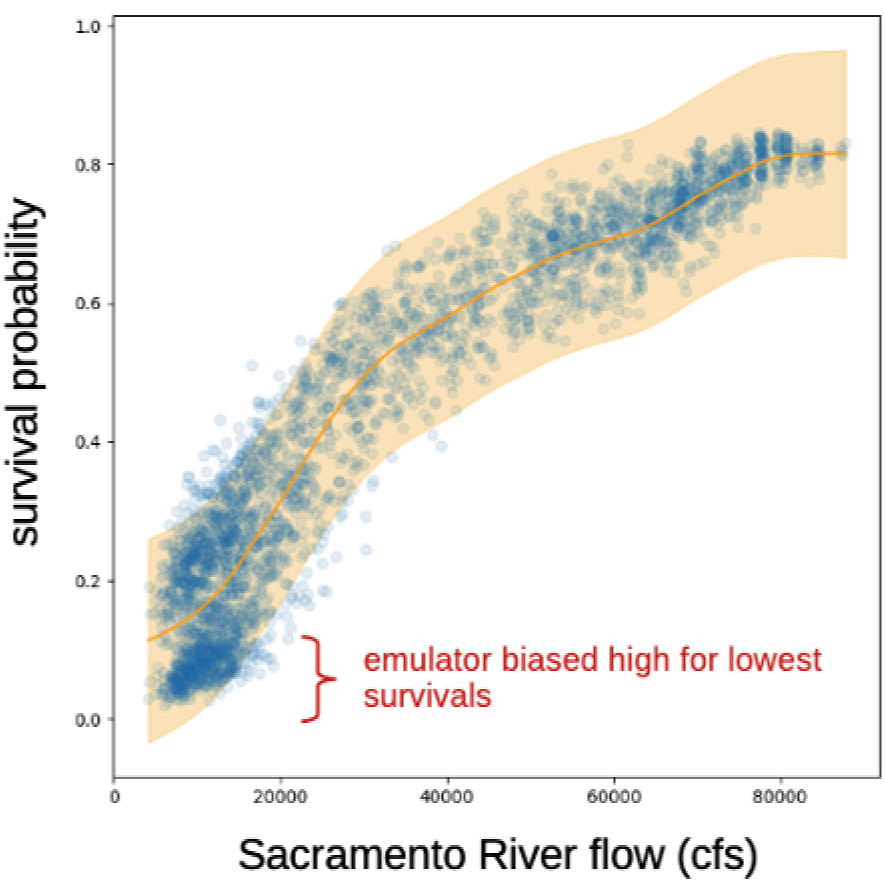


***Figure 9.*** *Simplified, one-variable emulator illustrating the root cause of the bias towards high survival at low survival probabilities*

Q-Q plots of the emulators' survival quantiles versus the ePTM's survival quantiles also show no systematic difference in the shapes of the distributions, other than the deviation at low survival probabilities explained above (Figure 10).

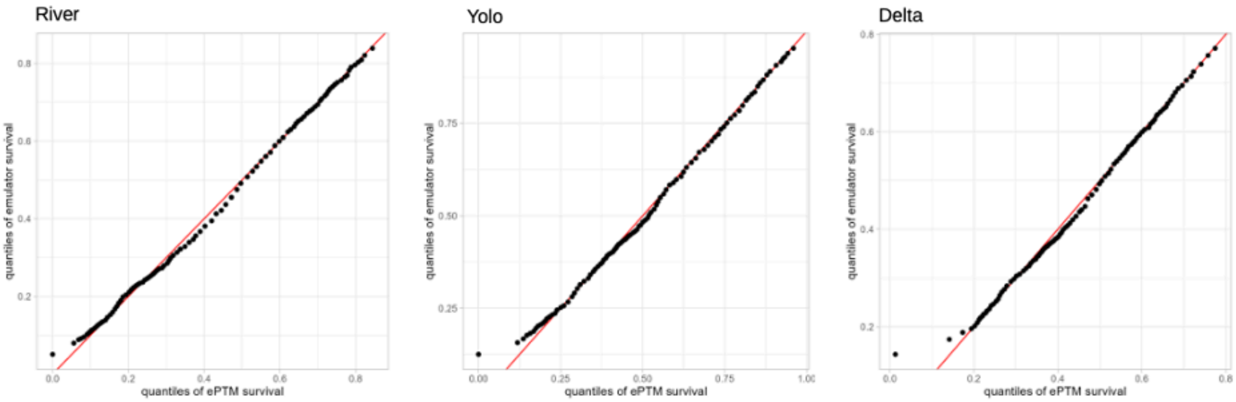


***Figure 10.*** *Q-Q plots of emulator survival quantiles versus ePTM survival quantiles*

## References

Gramacy, R. B. *Surrogates: Gaussian Process Modeling, Design, and Optimization for the Applied Sciences.* (Chapman and Hall/CRC, 2020).

Forrester, A. I. J., Keane, A. J. & Bressloff, N. W. Design and Analysis of 'Noisy' Computer Experiments. *AIAA J.* **44**, 2331–2339 (2006).

# Appendix S-3. Incorporating reintroduction to the McCloud River into the Winter-run Chinook salmon Lifecycle Model

Prior to the construction of the Shasta Dam, Chinook salmon spawned in the tributaries originating near Mount Shasta, including the McCloud River. Currently, there is interest in reintroducing Chinook to those spring-fed ancestral waters due the thermal refuge that such habitats can provide, particularly during the summer months. Multiple designs for the reintroduction are being proposed, such as volitional, semi-volitional, or trap and haul, and we propose a conceptual model that defines biological reintroduction rates common to all designs. The common rates are adult routing, adult survival, juvenile routing, and juvenile survival, which are required to complete the life-cycle from below Keswick to the McCloud and back.

We developed an additional sub-model for the WRLCM that includes reintroduction to the McCloud River. We made several assumptions about the spawning habitat quantity and egg incubation quality for winter-run Chinook in the McCloud River. Estimates of substrate for constructing redds on the McCloud river vary depending on whether 20m$^2$, 10m$^2$, or 6m$^2$ are used for redd sizes (USBR 2014). The highest estimate of redds assumed 6m$^2$ with a total estimate of 4,150 spawning sites of that size. As a result, we assumed that the capacity was 4000 female spawners in the McCloud for the purpose of this analysis. Thermal conditions for egg incubation are expected to remain below the critical threshold of 53.5 F (12 C) for much of the incubation period in the McCloud due to the river being spring-fed. There is the potential for the temperature on the McCloud to exceed the 53.5 F threshold in late July and early August, however (Anchor QEA and HDR 2026). As a result, the temperature dependent mortality was assumed to be 0.0 for April - July, and 0.05 in August.

The reintroduction sub-model is defined below as a series of steps that starts with the number of adult fish that transit from the Sacramento to the McCloud River and ends with the number of juveniles that enter the Sacramento River.

The steps to the reintroduction sub-model that are performed each year:

1. Set a target of female spawners for the McCloud of 2,000 female spawners. The proportion of fish reintroduced to the McCloud is calculated as the target of 2,000 relative to the total number of returning female spawners. The returning spawners are apportioned to the McCloud first, with the remaining spawning in the Sacramento (if there are additional fish beyond the target). The monthly spawning proportions, which are based on temperatures below Keswick, are used in both the Sacramento and the McCloud spawning.

2. Fish that are targeted for reintroduction to the McCloud, but not collected, spawn in the Sacramento. We assumed 0.95 of fish targeted for reintroduction are collected.

3. Fish that are reintroduced to the McCloud survive with a survival rate that is dependent on reintroduction quality. We assumed a 0.975 survival rate from the Sacramento River to the McCloud River to spawn.

4. Eggs are produced using a Beverton-Holt stock recruitment relationship for the McCloud and Sacramento. The McCloud annual capacity is assumed to be 4,000 female spawners and the Sacramento capacity is assumed to be 30,000 female spawners (Winship et al. 2014).

5. Fry emerge from the eggs according to the egg to fry survival relationship described in Osterback et al. (2026), which includes a temperature dependent egg to fry survival function. For the spawners in the Sacramento, the temperature below Keswick is used for calculating the temperature dependent mortality. In the McCloud River, the temperature dependent mortality is assumed to be 0.00 in June and July and 0.05 in August as described above.

6. Fry attempt to move out of the McCloud at a constant rate per month. Of those that chose to move down river, a proportion are successfully diverted to transit to the Sacramento, which is defined by the reintroduction quality. Fry that do not attempt to migrate or that are not successfully diverted in a given month remain in the McCloud until the following month with a McCloud fry survival rate. We assume that 70% of the fry attempt to move downriver in each month, and that 0.95 of them are able to be diverted successfully. The monthly fry survival in the McCloud is assumed to be 0.65, which is slightly higher than the fry survival rate in the Sacramento River of 0.50 due to the higher quality habitat there.

7. Fry that are successfully collected then move to the Sacramento River where they are released and join the juveniles in the Upper River habitat. The survival during transit to the Sacramento River is assumed to be 0.975.

The fry from the McCloud join the fry in the Upper River region of the WRLCM and are not tracked separately from this point forward in the life cycle model. They complete their life cycle by transitioning downstream through the additional habitats, smolting, entering the ocean, maturing, and returning to spawn as adults.

## References

Anchor QEA and HDR (HDR Engineering, Inc.), 2026. *Background Compendium and Design Criteria Report for the Feasibility of Volitional Fish Passage above Keswick and Shasta Dams*. Study of Physical Data Gaps to Inform the Implementation of Nur Rematriation Upstream of Shasta Dam (AB 211 Drought Grant Agreement Number – Q2396040). Prepared for California Department of Fish and Wildlife. January 2026.

USBR (United States Bureau of Reclamation) 2014. A Framework for Assessment of Habitat Conditions to Inform Planning for a Pilot-Level Chinook Salmon Reintroduction Study Shasta Dam Fish Passage Evaluation

https://www.usbr.gov/mp/bdo/docs/habitat-assessment-framework.pdf

Winship, Arliss J., Michael R. O ′Farrell, and Michael S. Mohr. 2014. Fishery and Hatchery Effects on an Endangered Salmon Population with Low Productivity. *Transactions of the American Fisheries Society*, no. 143: 957–71. https://doi.org/10.1080/00028487.2014.892532.