Appendix J, Winter and Spring Pulses and Delta Outflow **Attachment J.1 Longfin Smelt Outflow**

J.1.1 Model Overview

The potential effect of operations on Longfin Smelt abundance was investigated through development of a statistical modeling approach relating the Longfin Smelt Fall Midwater Trawl (FMWT) abundance index to: ([1](#page-0-0)) Delta outflow¹; (2) the FMWT abundance index two years earlier (as a representation of parental stock size), and; (3) ecological regime (i.e., 1967–1987, pre-*Potamocorbula amurensis* invasion; 1988–2002, post-*P. amurensis* invasion; and 2003–2022, Pelagic Organism Decline [POD]). The inclusion of the regime factor represents major ecological changepoints in the Bay-Delta (e.g., Nobriga and Rosenfield 2016; Sommer et al. 2007). Total Delta outflow (thousand acre-feet) was summed and examined as an explanatory covariate for two overlapping time periods: December through May, and March through May. Similar time periods have also been investigated in previous studies by Mount et al. (2013:66– 69) and Nobriga and Rosenfield (2016). Bayesian methods were used to account for model uncertainty (e.g., uncertainty in the time period over which Delta Outflow is considered to affect Longfin Smelt abundance), therefore integrating an important component of scientific uncertainty into the resulting model predictions for decision making.

J.1.2 Model Development

J.1.2.1 Methods

Twelve log-linear regression models were considered in the analysis. The models were fit to the FMWT index of Longfin Smelt abundance^{[2](#page-0-1)} (1967–2022) using a Bayesian approach implemented in the R statistical computing language (R Core Team 2023) via the *brms* package (Bürkner 2017). Three Markov Chain Monte Carlo chains were run for each model and flat priors were assumed for covariates. There was a 2,000-sample warm-up for each chain before 10,000 samples were retained as draws from the posterior (30,000 samples total drawn from the posterior). Bayesian values for the \hat{R} statistic were less than 1.01 across estimated parameters, which indicated sampling converged on the posterior probability distributions for all models considered.

¹ Downloaded from:<https://data.ca.gov/dataset/dayflow>

² Downloaded from:<https://apps.wildlife.ca.gov/FMWT>

Preliminary model comparison was performed using leave-one-out cross validation (LOO; Vehtari et al. 2017). Measures of model predictive accuracy using LOO are asymptotically equal to the widely applicable information criteria (WAIC; Watanabe 2010), but in the case of finite data LOO has been shown to be more robust to influential observations like outliers (Vehtari et al. 2017). The extent of model overlap in predictive accuracy was measured by the differences (and the standard errors of the differences) in expected log pointwise predictive densities, i.e., the differences in out-of-sample predictive accuracy between models. The preliminary model comparisons indicated there was a relatively high degree of similarity in terms of predictive ability between the top scoring individual models.

Therefore, rather than selecting a single model for inference, the posterior predictive probability distributions were combined as a weighted average across models. This process involved taking draws from the posterior of each single model in proportion to its model weight, with model weights for averaging posterior predictive distributions calculated using the *loo* package (Vehtari et al. 2020). For example, if a single model's weight was 25 percent of the total model set, then 2,500 draws from its posterior were added to the averaged posterior predictive distribution, which included 10,000 total draws taken across the posterior predictive distributions for all models. The statistical approach used to calculate the model weights for averaging the posterior predictive distributions across models is known as "stacking" (Yao et al. 2018).

Compared to more traditional model averaging approaches, stacking differs in terms of how model weights are assigned. Instead of calculating model weights based on the relative predictive ability for each individual model—where the best model for prediction would be given the highest weight—the model weights estimated through stacking minimize the LOO mean squared error of the resulting averaged posterior predictive distribution across models. In other words, stacking was used to estimate the optimal linear combination of model weights for averaging predictive distributions across the model set (Yao et al. 2018).

Hence, the model with the largest stacking weight does not necessarily have the highest predictive score compared to other models in the set. For example, the models in this case can be divided into two subsets: one subset includes a covariate for Delta outflow during December-May and the other model subset includes a covariate for March-May Delta outflow (Table J.1-1). Comparing the predictive ability of each individual model using LOO resulted in a model with December-May outflow (the model with the third highest stacking weight in Table J.1-1) having the highest individual predictive accuracy of any single model considered. In contrast, when the optimal linear combination of weighted model predictions was calculated, stacking resulted in a model with March–May Delta outflow having the highest single model weight (37 percent of the total stacking weight across the model set). Nevertheless, because stacking optimizes the linear combination of model weights for predictive accuracy, the next four models (~63 percent of the stacking weight) all include December–May Delta outflow instead of March–May Delta outflow. Therefore, in this case, even though the model with highest stacking weight included March– May Delta outflow, the averaged posterior predictive distribution was ultimately weighted more heavily with models that include December–May Delta outflow compared to models with March–May Delta outflow. Of the twelve models considered, the top five models by stacking weight accounted for >99.9 percent of the averaged posterior predictive distribution (Table J.1- 1).

Predictions of the fall midwater trawl abundance index under the modeled CalSim 3 outflow scenarios (1922–2022) were generated using the model stacking approach described above to generate a weighted average Bayesian posterior predictive distribution across the set of models considered. Dropping subscripts denoting individual models for simplicity, the general form of the models can be written as:

$$
Log_{10}[FMWT_{yr}] \sim N(\mu_{yr}, \sigma^2)
$$
 (1)

$$
\mu_{yr} = \beta_{0,i} + \beta_1 Outflow_{yr,j} + \beta_2 Log_{10}[FMWT_{yr-2}] + \beta_3 Regime_i * Outflow_{yr,j} \tag{2}
$$

where:

- $Log_{10}[FMWT_{yr}]$ is the model predicted Log_{10} value of the fall midwater trawl index in water year *yr*;
- \bullet μ_{vr} is the expected fall midwater trawl index in water year *yr* (the stacked posterior predictive distribution for μ_{vr} is shown as the dark grey ribbon in Figure J.1-1 Figure J.1-1);
- \bullet σ^2 is the residual variance parameter (the stacked posterior predictive distribution including the residual variance is shown as the light grey ribbon in [Figure J.1-1F](#page-3-0)igure J.1-1);
- \bullet $\beta_{0,i}$ represents the intercept parameter estimated for each regime: Pre-*Potamocorbula (i* = 1); *Potamocorbula (i* = 2); and POD ($i\beta_0$ ^[OBJ]; $\beta_{0,i}$ ^[OBJ];
- β_1 represents the slope parameter estimated for the relationship between the fall midwater trawl index and Delta outflow;
- *Outflow_{yr,j}* is the normalized^{[3](#page-2-0)} outflow level during water year *yr*, and *j* denotes the outflow level during either the December through May, or the March through May period;
- β_2 represents the slope parameter estimated for the relationship between the expected fall midwater trawl index and the value of that index 2 years prior. For models without the parental stock covariate, $\beta_2 = 0$, and;
- β_3 represents the interaction covariate (the difference in slopes) with respect to the estimated effect of outflow on the FMWT index of abundance during different regimes (The asterisk "*" sign represents an interaction term between Regime and Delta Outflow). For models without this interaction term, $\beta_3 = 0$.

³ Normalized outflow values for each CalSim 3 scenario were calculated by subtracting the mean and dividing by the standard deviation of observed Delta outflow values (1967–2020).

Note: The circles represent the annual historical values of the fall midwater trawl abundance index. Colors correspond to the three modeled regimes. The solid lines connect the annual expected values from the stacked Bayesian posterior distribution. The darker gray ribbon represents the 95% posterior probability interval for the annual expected values for the fall midwater trawl index value. The lighter gray ribbon with a dashed black outline represents the averaged 95% overall posterior predictive probability interval for the fall midwater trawl index value. The posterior uncertainty interval (dark gray ribbon) has a smaller range than the posterior predictive interval (light gray ribbon) because in addition to uncertainty in the estimated model covariate values, the posterior predictive distribution also incorporates uncertainty in the residual error of the model fits (Equations 1 and 2 below).

Figure J.1-1. Stacked Posterior Predictive Distributions for the Log-Linear Regressions of Longfin Smelt Fall Midwater Trawl Abundance Index as a Function of Delta Outflow (December–May), Ecological Regime (1967–1987, pre-*Potamocorbula amurensis* invasion; 1988–2002, post-*Potamocorbula* invasion [shown as *Potamocorbula*]; and 2003–2022, Pelagic Organism Decline [POD]), and Abundance Index 2 Years Earlier $[Log₁₀ FMWT(yr - 2)]$.

For those models that included the $Log_{10} FMWT(yr-2)$ parental stock size covariate (Table [J.11T](#page-4-0)able J.1-1), the starting parental stock size in 1922 and 1923 was set at a FMWT index value of 118.2, corresponding to the mean index value from 2013 through 2022. Given the starting values for the FMWT index (in the relevant models), the recursive nature of the regression formula was used to generate the expected FMWT index value in successive years from the posterior predictive distribution two years prior. For all models, predictions were conditional on the estimated relationship between the FMWT index and Delta outflow (in December–May, or March–May, depending on the model), and for those models that included a regime covariate, draws from the posterior predictive distributions were conditioned on estimates during the Pelagic Organism Decline regime.

Table J.111-1. The Optimal Linear Combination of Model Weights based on Stacking, which Minimizes the Mean Squared Error of the Leave-One-Out Cross Validation for the Resulting Model Averaged Posterior Predictive Distribution across the Twelve Log-Linear Regressions of Longfin Smelt Fall Midwater Trawl Abundance Index. Models are a Function of Delta Outflow (December–May or March–May), Ecological Regime (1967– 1987, pre-*Potamocorbula amurensis* invasion; 1988–2002, post-*P. amurensis* invasion; and 2003–2022, Pelagic Organism Decline), and Abundance Index 2 Years Earlier (Log₁₀ $FMWT(vr - 2)$).

^a An asterisk "*" sign represents an interaction term between Regime and Delta Outflow.

As an example, starting in 1924, draws from the posterior predictive distribution for models including the parental stock size covariate were generated by first substituting the normalized 1924 December through May (or March through May) CalSim 3 outflow value for each alternative. CalSim 3 outflow values prior to normalization (i.e., in units of million acre-feet) are shown in Figure J.1-2. Draws from the posterior distributions for the regression parameters and

the starting value for $Log_{10}[FMWT_{1922}]$ were then used to generate the posterior predictive distribution for the fall midwater trawl index in 1924 (μ_{1924}). This value was then substituted into Equation 1, and the posterior distribution for the residual variance parameter was used to generate draws from the pointwise posterior predictive distributions for the fall midwater trawl index[.](#page-5-0)⁴ This process was iterated over each successive year, substituting the derived μ_{yr-2} values for $Log_{10}[FMWT_{vr-2}]$ to calculate μ_{vr} , and to generate the annual posterior predictive distributions for the fall midwater trawl index under each alternative. For models that did not include the parental stock size covariate, the posterior predictive distributions were generated based on the corresponding CalSim 3 outflow values for the monthly period corresponding to the individual model estimates, and likewise conditioned on covariate estimates during the POD regime for models that included a regime covariate (or the constant intercept parameter β_0 , for models without the regime covariate). As noted above in the description of the model stacking approach, draws from the posterior predictive distribution for each model were sampled in proportion to the stacking model weights, to generate a weighted average posterior predictive distribution across the models considered. Summaries were then calculated by grouping the stacked annual posterior predictive distributions by water year type and calculating the means and Bayesian credible intervals for each aggregated water year type posterior predictive distribution.

J.1.2.2 Assumptions / Uncertainty

Several additional models were also examined, in addition to those in [Table J.11T](#page-4-0)able J.1-1, but they were ultimately not included in this analysis due to poor model fits and what would have been additional computational cost without an expected difference in results (i.e., the poor model fits are indicative of poor model predictive accuracy, and hence tiny model weights). The additional models included a squared term on Delta outflow and their examination was motivated by the modeling results of Nobriga and Rosenfield (2016). Those authors assessed the relationship between Delta outflow and the ratio of age-0 to age-2 Longfin Smelt abundance in the two-life-stage versions of the models included in their analyses. They found support for nonlinearity in this relationship (i.e., there was a peak in productivity at more intermediate outflow values), which led to the inclusion of a second-order polynomial regression (i.e., a squared term) on Delta outflow (Nobriga and Rosenfield 2016:50). Given the approach taken here, which differs from the Nobriga and Rosenfield analysis in terms of: (1) the survey data used for Longfin Smelt abundance; (2) how Delta outflow values were included as covariates, and; (3) the overall time periods for available data included in the regression models, there was little to no support found for a second-order polynomial regression on Delta outflow. The aforementioned factors that differed between the two analyses are briefly described in the next paragraph for completeness; but, given the poor predictive ability of the second-order polynomial regressions under the current approach, that subset of models was ultimately not included because the preliminary results indicated the stacked model weights would be near zero. Hence the averaged posterior predictive distributions would not be expected to be sensitive to the exclusion of those models in this case, but their inclusion would have increased the computational time necessary to run and perform the averaging over a larger set of models.

⁴ "~*N*" in Eqn. 1 denotes a normal (Gaussian) distribution.

As outlined above, there are several differences between these analyses and those of Nobriga and Rosenfield (2016) that might explain the discrepancy in terms of support (or lack thereof) found for dome shaped Longfin Smelt productivity as a function of Delta outflow. Firstly, Nobriga and Rosenfield (2016) found support for this relationship fitting models to catch data from the San Francisco Bay Study. In these analyses, on the other hand, the regression models have been fit to the FMWT index of abundance instead. Second, Nobriga and Rosenfield (2016) incorporated covariate values for Delta Outflow based on a principal component analysis (the first principal component values) of the *z*-scored monthly means from December to May. Here, the monthly total outflows (either from December to May, or March to May) were summed, resulting in a total outflow value during each time period each year, and the regression covariate values were calculated as the *z*-scores of the period-total outflow values taken across years. Third, in addition to examining indices of abundance from different surveys, the annual time periods that have been examined also differ. Nobriga and Rosenfield (2016) examined the relationship between annual indices of Longfin Smelt abundance-at-age and Delta outflow that were available from the Bay Study during 1980–2013. In these analyses, this relationship was examined over a longer period (1967–2022), which includes >20 additional years in the comparison between Longfin Smelt abundance and Delta outflow.

This evaluation of management scenarios assumes that the correlation between outflow and the FMWT index of Longfin Smelt abundance during the POD regime (2003-2022) will remain constant in the future. While a positive correlation between outflow and the FMWT index has been evident for over 50 years of FMWT surveys, the strength of this relationship does not appear constant through time. For example, the results of Bayesian stacking indicates that averaging over models with ecological regime (at least as defined here) maximizes the out-ofsample predictive ability of models. In other words, these results indicate the FMWT data are consistent with shifts in stock productivity between regimes as a function of outflow. Only one model, with no regime covariate (equivalent to assuming the FMWT correlation with outflow has been constant since 1967) was assigned any substantial $(\sim 10\%)$ weighting in the final averaged model set (Table J.1-1). In the absence of empirical data on whether or how this correlation may change in the future, it is reasonable to evaluate the management scenarios based on the correlation during the most recent ecological regime.

J.1.2.3 Code and Data Repository

Analysis files and code for the Longfin Smelt Outflow analysis are available upon request.

J.1.3 Results

Table J.11-22. Means of annual posterior predictive means for the FMWT index of Longfin Smelt abundance by water year type (WYT) for BA scenarios.

Table J.1-313. Means of annual posterior predictive means for the FMWT index of Longfin Smelt abundance by water year type (WYT) for EIS Scenarios. The percentage difference between scenarios and NAA is shown in the parentheses.

Table J.1-414. Means of annual posterior predictive distributions for the FMWT index of Longfin Smelt abundance. Water year types (WYT) are shown by first initial (see [Table](#page-4-0) [J.11T](#page-4-0)able J.1-1 to reference full names for each type). The percentage difference between scenarios and NAA is shown in the parentheses.

Table J.1-5. 15Means of annual posterior predictive distributions for the FMWT index of Longfin Smelt abundance. Scenarios considered in the BA are shown for comparison across years.

Table J.1-61. The percentages of each water year type are shown for: (1) the CalSim 3 scenarios, and; (2) the historical water year types during 2003-2022 (POD regime).

Figure J.1-212. CalSim 3 outflow values are shown for comparison across BA scenarios.

Figure J.1-3. Posterior predictive distributions for the FMWT index of Longfin Smelt abundanc[e](#page-18-0) are shown aggregated by water year type⁵ for each BA scenario. The horizontal line in the distribution for each scenario represents the median predicted value.

⁵ https://cdec.water.ca.gov/reportapp/javareports?name=WSIHIST

Figure J.1-4. Posterior predictive distributions for the FMWT index of Longfin Smelt abundance are shown aggregated by water year type for each EIS scenario. The horizontal line in the distribution for each scenario represents the median predicted value.

Figure J.1-5. The 95th Bayesian credible intervals for the posterior predictive distributions are shown, based on the parental stock model and the 100 year time series of CalSim 3 Delta Outflow values for each BA scenario.

Figure J.1-6. The 95th Bayesian credible intervals for the posterior predictive distributions are shown, based on the parental stock model and the 100 year time series of CalSim 3 Delta Outflow values for each EIS scenario. The credible intervals for the NAA scenario are overlaid as the dashed black lines for comparison with the alternatives.

Figure J.1-7. Posterior predictive distributions for the FMWT index are shown as a function of hypothetical outflow levels. The hypothetical outflow levels are paired between the two monthly time periods considered. The paired (Dec-May / Mar-May) values are based on the linear correlation of cumulative outflow values generated from the CalSim 3 runs. "MAF" denotes cumulative outflow across the monthly time periods in million acre-feet.

J.1.4 References

- Bürkner, P.-C. 2017. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* 80(1):1–28.
- Mount, J., W. Fleenor, B. Gray, B. Herbold, and W. Kimmerer. 2013. *Panel Review of the draft Bay-Delta Conservation Plan.* Prepared for the Nature Conservancy and American Rivers. September. Saracino & Mount, LLC, Sacramento, CA.
- Nobriga, M. L., and J. A. Rosenfield. 2016. Population Dynamics of an Estuarine Forage Fish: Disaggregating Forces Driving Long-Term Decline of Longfin Smelt in California's San Francisco Estuary. *Transactions of the American Fisheries Society* 145(1):44–58.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. [https://www.R-project.org/.](https://www.r-project.org/)
- Sommer, T., C. Armor, R. Baxter, R. Breuer, L. Brown, M. Chotkowski, S. Culberson, F. Feyrer, M. Gingras, B. Herbold, W. Kimmerer, A. Mueller-Solger, M. Nobriga, and K. Souza. 2007. The collapse of pelagic fishes in the upper San Francisco Estuary. *Fisheries* 32(6):270–277.
- Vehtari, A., J. Gabry, M. Magnusson, Y. Yao, P. Bürkner, T. Paananen, and A. Gelman. 2020. *loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models.* R package version 2.4.1. Available: [https://mc-stan.org/loo/.](https://mc-stan.org/loo/) Accessed: May 11, 2022.
- Vehtari, A., A. Gelman, and J. Gabry, 2017. Practical Bayesian model evaluation using leaveone-out cross-validation and WAIC. *Statistics and Computing* 27(5):1413–1432.
- Watanabe, S. 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research* 11:3571–3594.
- Yao, Y., A. Vehtari, D. Simpson, and A. Gelman. 2018. Using Stacking to Average Bayesian Predictive Distributions (with Discussion). *Bayesian Analysis* 13(3):917–1007.

This page intentionally left blank.