Patterns of Synchrony and Environmental Thresholds in the Performance of Forecast Models Used for U.S. West Coast Chinook and Coho Salmon Stocks

William H. Satterthwaite¹, Kelly S. Andrews², Jennifer L. Gosselin³, Correigh M. Greene⁴, Chris J. Harvey⁵, Mary Hunsicker⁵, Stuart H. Munsch⁶, and Jameal F. Samhouri²

¹Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic & Atmospheric Administration, 110 McAllister Way, Santa Cruz, California 95060, USA
²Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic & Atmospheric Administration, Seattle, Washington 98112, USA
³School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98105, USA
⁴Fish Ecology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic & Atmospheric Administration, Seattle, Washington 98112, USA
⁵Fish Ecology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic & Atmospheric Administration, Newport, OR 97365, USA
⁶Ocean Associates Inc., Under Contract to Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic & Atmospheric Administration, 2725 Montlake Boulevard East, Seattle, Washington 98112, USA

Keywords: nonlinear, ecosystem-based management, environmental indicator, portfolio, threshold

Preseason abundance forecasts drive management of salmon fisheries off the U.S. West Coast, yet little is known about how environmental variability influences forecast performance. Understanding shared patterns in forecast performance can identify scenarios with heightened management risk due to shared over/under-forecasting of a large proportion of the fishery stock portfolio, and identifying drivers of this synchrony may aid in the development of improved forecasts. We examined temporal patterns, synchrony, and potential drivers in shared trends of forecast performance for 21 Chinook and 15 coho salmon stocks on the U.S. West Coast. For select Chinook salmon stocks of particularly high management importance, we tested for nonlinear and threshold relationships between forecast performance and environmental indices.

Fig. 1. Heat map of log (postseason estimate/forecast) by year (white = no data). Note that within each species, stocks are ordered as they would be encountered moving along the coastline from north to south. Occasional periods of nearly uniform under- or over-forecasting across stocks seem visually apparent within species, especially for coho. Sustained runs with multiple consecutive years of under- or over-forecasting within stocks appear more common for Chinook.
We quantified correlation and synchrony in stock forecast performance (Fig. 1, defined as log of postseason abundance estimates divided by preseason forecasts) using data from a core set of years 1997–2016 for which we had forecast performance data for all stocks except Lower Columbia Natural Coho and Upriver Columbia Summer Chinook (which we excluded from this portion of the analysis). For straightforward interpretation, we calculated the mean pairwise correlation across stocks for all stocks combined, for all Chinook salmon stocks, and for all coho stocks. To further explore potential common trends in forecast performance, we used dynamic factor analysis (DFA), implemented via the MARSS R package (Holmes et al. 2012), to identify common trends in forecast performance and the loading of each stock onto these trends. Because MARSS can accommodate missing values, we included all available data in this analysis. We tested whether stocks tended to cluster together in their factor loading due to geography, forecast type, hatchery influence, or other factors. We also examined correlations between shared trends extracted by the DFA and environmental indices.

Fig. 2. Factor loadings for the best-supported dynamic factor analysis model applied to coho stocks. The inset shows the shared trends.

Forecast performance was asynchronous across all stocks and species (mean pairwise correlation of $r = 0.10$) but slightly more synchronous within species ($r = 0.14$ for Chinook, $r = 0.23$ for coho). Most strong positive correlations in forecast performance were between geographically proximate stocks. DFA applied to all stocks revealed a single shared trend, seen most strongly among southern coho stocks. However, the best-supported DFA model explained only 13% of the variance and loading on the single shared trend was low for most stocks. The best-supported DFA model for Chinook extracted only one trend and explained only 18% of the variance. For coho, the best-supported DFA model consisted of two shared trends, explaining 32% of the variance (Fig. 2). Loadings for coho stocks tended to cluster geographically but not with respect to hatchery versus wild nor by forecast type.

Our exploration of nonlinearities and thresholds focused on Chinook stocks of particular concern for United States West Coast fisheries management and conservation: Klamath and Sacramento River fall Chinook which are key ocean fishery stocks recently declared overfished, and Puget Sound Chinook stocks which were identified as the highest priority prey for endangered Southern Resident Killer Whales. We tested for nonlinearities and thresholds in the relationships between putative environmental drivers (freshwater conditions during spawning and rearing, localized ocean conditions at ocean entry, and intermediate to basin-scale ocean conditions throughout ocean residency) and forecast performance by comparing linear and nonlinear (generalized additive models, GAMs) models. We considered 95% confidence intervals on the second derivative of the fitted relationship excluding zero as evidence for a threshold (Large et al. 2013).

For this analysis, we calculated annual forecast performance $P_y$ for each stock as:

$$P_y = \frac{f_y - o_y}{\frac{1}{N} \sum_{i=y_{\text{min}}}^{y_{\text{max}}} f_i - o_i}$$

where $N$ is the number of years with data, $f_i$ is the preseason forecast and $o_i$ is the postseason observation or estimate for year $y$. Positive values indicate that fewer Chinook returned to spawn than expected (overforecasting),
negative values indicate more Chinook returned to spawn than predicted (underforecasting), and values far from zero indicate unusually large errors.

Because our exploratory approach tested a large number of stock-index-location-lag combinations, spurious relationships were a concern. We therefore simulated 200 versions of each forecast performance timeseries by randomly resampling (with replacement) a score for each year, modeled the relationship between the resampled timeseries and the environmental indices tested for the corresponding stocks, and tracked the frequency of simulations where a nonlinear model was selected as well as the distribution of $R^2$ values. Given a proportion $s$ of resampled relationships with $R^2$ above a critical threshold $C$, and $k$ fitted relationships with $R^2 > C$ in the empirical data in $n$ stock-index-location-lag test combinations, we calculated the probability of observing at least $k$ relationships at least this strong by chance using a Bernoulli model.

We found 13 cases (12 nonlinear) where an environmental index could explain at least 50% of the variation in forecast performance and 55 cases (42 nonlinear) where the index could explain >33% of the variation in forecast performance. No relationships with $R^2 > 0.16$ were found for Klamath River Fall Chinook while two relationships with $R^2 > 0.40$ were found for Sacramento River Fall Chinook. All other relationships with $R^2 > 0.33$ were for Puget Sound stocks. The null model suggests it is unlikely we would see so many cases of $R^2 > 0.33$ by chance ($p = 0.16$ for all Puget Sound stocks combined, $p = 0.0012$ for South Puget Sound natural summer-fall Chinook) but the number of relationships with $R^2 > 0.50$ observed is consistent with null model expectations. However, comparing against the null model is likely conservative because tests of forecasts based on different approaches arguably reflect distinct hypotheses, and we did not consider all driver-lag-model combinations equally likely a priori (e.g., we hypothesized sibling-based models would be most sensitive to recent ocean conditions and relatively insensitive to freshwater or long lags).

For Sacramento Fall Chinook, which uses a sibling-based forecast, the top two environmental indices explaining errors in preseason forecasts were related to ocean conditions in the year of return (Pacific Decadal Oscillation (PDO) and the North Pacific Index). This is concurrent with when managers need to make forecasts, but PDO can be predicted one year in advance with some skill (Lienert and Doblas-Reyes 2013). In addition, it is consistent with the expectation that performance of sibling-based forecasts would be most affected by conditions experienced after the return of younger members of the cohort that inform the forecasts. For Puget Sound stocks, which employ a variety of forecasting methods, a variety of indices operating over a range of lags displayed good explanatory power, although overall freshwater indices were rarely supported, possibly because freshwater effects are directly or indirectly incorporated (e.g., via smolt counts) into the forecasts.

We found evidence of thresholds in most (60/65) cases where nonlinear models were preferred. Figure 3 displays an illustrative relationship between Sacramento Fall Chinook forecast performance and PDO in the year of return, which took on extreme values in 2008–2009, years associated with a fishery collapse and closure. Returns of many Puget Sound stocks seemed to show a shared response in overforecasting 2014 returns, which was correlated with unusual sea level height off Alaska in 2013. Further research into conditions characterizing the 2014 return year is advised.

The individual relationships identified here should be approached with caution due to the exploratory nature of this study, but warrant further investigation and consideration by managers. Null model results suggest that individual relationships should be approached with caution, but it is unlikely that they are all spurious. The thresholds we identified here identify conditions under which precautionary management may be warranted in a
particular year, suggest that some indices merit consideration for inclusion in forecasts, and offer insights into ways forward for improving salmon forecasts given increasingly dynamic ocean conditions.

REFERENCES

Holmes, E., E. Ward, and K. Wills. 2012. MARSS: Multivariate Autoregressive State-space Models for analyzing time-series data. The R Journal. 4: 11–19.

Large, S.I., G. Fay, K.D. Friedland, and J.S. Link. 2013. Defining trends and thresholds in responses of ecological indicators to fishing and environmental pressures. ICES J. Mar. Sci. 70: 755–767.

Lienert, F. and F.J. Doblas-Reyes. 2013. Decadal prediction of interannual tropical and North Pacific sea surface temperature. J. Geophys. Res.: Atmospheres 118: 5913–5922.