Eulachon (*Thaleichthys pacificus*) marine ecology: applying ocean ecosystem indicators from salmon to develop a multi-year model of freshwater abundance

Sarah A. Montgomery

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Marine Affairs

University of Washington

2020

Committee:

Cleo Woelfle-Erskine

Robert Anderson

Program Authorized to Offer Degree:

Marine and Environmental Affairs

Abstract

Eulachon (*Thaleichthys pacificus*) marine ecology: applying ocean ecosystem indicators from salmon to develop a multi-year model of freshwater abundance

Sarah A. Montgomery

Chair of the Supervisory Committee:
Dr. Cleo Woelfle-Erskine
School of Marine and Environmental Affairs

Complex ecological processes determine whether and how fish species survive in the ocean. For some fishes such as Pacific salmon (*Oncorhynchus* spp.), these processes are thoroughly studied and modeled. For eulachon (*Thaleichthys pacificus)*, however, these processes are more mysterious despite their trophic link to salmon and similar anadromous life history. The ocean ecology of eulachon was identified by the National Oceanic and Atmospheric Administration (NOAA) as a key knowledge gap and research priority for eulachon, which were listed under the Endangered Species Act as a threatened species in 2010 after suffering significant declines throughout the southern portion of their range. Eulachon are a culturally and historically important species for Native nations, as well as a critical component of freshwater, estuarine, and marine food webs. So, a decline in their abundance has motivated researchers to better understand what drives fluctuations in their populations, despite a lack of data in the marine environment.

The marine ecology of eulachon may be relatively understudied, but data collected and compiled for salmon can partially fill this gap. NOAA developed ocean ecosystem indicators (physical, chemical, and biological factors) that are used to predict salmon returns. I applied these indicators to a new question: are the indicators used for salmon also predictive for a trophically related and ecologically similar species? And what environmental and biological factors in the ocean drive fluctuations in eulachon abundance in a major spawning basin, and when?

Using multivariate analyses, I found that ocean ecosystem indicators in years of ocean residency are correlated with eulachon abundance in the Columbia River. Large-scale and bottom-up indicators such as the status of the Pacific Decadal Oscillation and prey abundance describe much of the variation in eulachon abundance. Time series analysis also indicates eulachon abundance correlates strongly with ocean conditions in the two and three years prior to their return, suggesting dominant life histories of two- and three-year ocean types. These results are promising for future modeling efforts for eulachon—ocean ecosystem indicators can be used to understand variability in eulachon populations and can inform recovery decisions and actions.

Eulachon (*Thaleichthys pacificus*) marine ecology: applying ocean ecosystem indicators from salmon to develop a multi-year model of freshwater abundance

Abstract

Complex ecological processes determine whether and how fish species survive in the ocean. For some fishes such as Pacific salmon (*Oncorhynchus spp.*), these processes are thoroughly studied and modeled. For eulachon (*Thaleichthys pacificus)*, however, these processes are more mysterious despite their trophic link to salmon and similar anadromous life history. The ocean ecology of eulachon was identified by the National Oceanic and Atmospheric Administration (NOAA) as a key knowledge gap and research priority for eulachon, which were listed under the Endangered Species Act as a threatened species in 2010 after suffering significant declines throughout the southern portion of their range. Eulachon are a culturally and historically important species for Native nations, as well as a critical component of freshwater, estuarine, and marine food webs. So, a decline in their abundance has motivated researchers to better understand what drives fluctuations in their populations, despite a lack of data in the marine environment.

The marine ecology of eulachon may be relatively understudied, but data collected and compiled for salmon can partially fill this gap. NOAA developed ocean ecosystem indicators (physical, chemical, and biological factors) that are used to predict salmon returns. I applied these indicators to a new question: are the indicators used for salmon also predictive for a trophically related and ecologically similar species? And what environmental and biological factors in the ocean drive fluctuations in eulachon abundance in a major spawning basin, and when?

Using multivariate analyses, I found that ocean ecosystem indicators in years of ocean residency are correlated with eulachon abundance in the Columbia River. Large-scale and bottom-up indicators such as the status of the Pacific Decadal Oscillation and prey abundance describe much of the variation in eulachon abundance. Time series analysis also indicates eulachon abundance correlates strongly with ocean conditions in the two and three years prior to their return, suggesting dominant life histories of two- and three-year ocean types. These results are promising for future modeling efforts for eulachon—ocean ecosystem indicators can be used to understand variability in eulachon populations and can inform recovery decisions and actions.

Introduction

Thaleichthys pacificus means fatty fish of the Pacific. Eulachon, also called candlefish, savior fish, and many other names, are culturally important to Native nations of the Pacific coast for their nutritional content and life history; eulachon return soon after the winter when other food fish are absent (Leland and Mitchell, 2001). Like many other fishes of the Pacific coast, populations of eulachon have declined due to environmental conditions in freshwater and marine habitats, commercial fisheries (bycatch, in the case of eulachon), dams and water diversions, water quality, and other factors. Eulachon have declined markedly since 1994, and the Southern Distinct Population Segment (DPS) was listed by the National Oceanic and Atmospheric Administration (NOAA) as "threatened" under the Endangered Species Act in 2010 (FR 75:13012; NOAA, 2017). Scientists are motivated to understand the variability and decline in eulachon populations because in addition to being culturally important to tribal and non-Native fishers as harvest, eulachon are a critical component of freshwater, estuary, and marine food webs upon which many other species rely. As a forage fish, eulachon consume energy (e.g., copepods) from lower trophic levels and make it available to predators such as salmon, piscivorous birds, and marine mammals (Peterson et al., 2014).

This study builds on and contributes to work in eulachon ecology and trophic dynamics in the California Current upwelling system. Some researchers have examined eulachon ecology in freshwater and estuarine systems (Litz et al., 2014; Mallette, 2014), and others have examined how ocean conditions affect the survival and abundance of other forage fishes and salmonids (Burke et al., 2015; Peterson et al., 2014), but few studies analyze how ocean conditions drive eulachon abundance. To fill this gap, this study provides additional insight into the ocean ecology of eulachon and specific biological, physical, and chemical factors in the ocean that may drive their abundance. Though numerous researchers (Gustafson et al., 2010; NMFS, 2016) have identified ocean conditions as a primary threat to eulachon recovery, the relationship between ocean conditions and abundance in a major spawning river has not been described. I address this research gap by demonstrating that ocean ecosystem indicators developed for salmon are good predictors of eulachon abundance in the Columbia River. This study also demonstrates how statistical methodologies can harness relevant, related data to understand an understudied species. The analytical focus on using available data for a closely related species is another contribution to fisheries science and recovery processes.

Eulachon are a semelparous forage fish in the family Osmeridae. They spawn in freshwater then die. Eggs incubate in sediment for around 30-40 days and wash out to estuaries with spring freshets (NMFS, 2017). The juveniles rear in estuaries then outmigrate to the ocean, where they spend 2-5 years. Eulachon return to freshwater in late winter through spring to spawn (NMFS, 2017). Eulachon spawn from northern California to southwestern Alaska, with the Southern DPS occupying rivers from central British Columbia to the Mad River in California (Figure 1). Within the Southern Distinct Population Segment, the Fraser and Columbia river basins produce the most eulachon (Gustafson et al., 2010). Since the Southern DPS declined in the mid-1990s, agencies in Canada and the U.S. have monitored and estimated run sizes in both rivers (NMFS, 2017). The collection of data in freshwater backlights knowledge gaps in their marine abundance and distribution.

The familiar anadromous narrative of spawning, outmigrating, rearing, and returning to freshwater glosses over a key eulachon life stage—their time in the ocean. Eulachon are estimated to spend 95-98% of their lives in the ocean but little information is available about their behavior and physiology during this time (Hay and McCarter, 2000, cited in NMFS, 2017). What is known is that eulachon populations display decadal to multi-decadal cycles that do not appear to synchronize with known changes in freshwater conditions. Another confusing piece of their life history is that abundance in any river system can vary highly year-to-year (Gustafson et al., 2010). Eulachon are known for their highly plastic life history characteristics, with variability in spawning dates and in age class structure within and between major river basins (NMFS, 2017). Both the decadal to multi-decadal variation and the year-to-year variation are puzzles that oceanic data may help answer. One theory is that eulachon recruitment in North Pacific marine ecosystems responds to shifts in ocean-atmosphere conditions (Gustafson et al., 2010; NMFS, 2017), but the precise interactions between eulachon and the physical, chemical, and biological processes in the ocean are unknown. This unknown is a problem to recovering eulachon.

Part of NOAA's recovery planning process involves identifying threats to species recovery. In 2010, NOAA's Biological Review Team identified sixteen threats to eulachon "that alter key physical/biological and/or chemical features and reduce a species' viability" (Gustafson, 2010; NMFS, 2017). Climate change impacts on ocean conditions was the only threat ranked as high severity in every subpopulation (Klamath, Columbia, Fraser, and British Columbia coastal rivers). The known and potential mechanisms for effects of climate change are described in the Recovery Plan (NMFS, 2017), and the subpopulations represent major spawning aggregates or river basins within the Southern DPS, as genetic analyses have not elucidated population structure at a smaller scale than the DPS. Other threats vary in their severity geographically and by subpopulation. Some of the other most severe threats include climate change impacts on freshwater habitat, dams/water diversions, bycatch, and predation. Because the impacts of climate change on ocean conditions are considered not only of high severity but also highly uncertain due to the lack of information about how eulachon populations respond to ocean conditions, the relative impact of changing ocean conditions on eulachon recovery received the highest ranking and much discussion in the Recovery Plan. The lack of data exacerbates the threat. To address this threat, the Recovery Plan identifies the following research priority related to Species Ecosystem Interactions:

> "Conduct a gap analysis to identify the data needs to develop an ocean ecosystem indicators model of eulachon marine survival in the California Current Ecosystem [then,] develop an ocean ecosystem indicators model of eulachon marine survival in the California Current Ecosystem to determine how short-term and long-term variability in ocean conditions affect eulachon abundance and productivity for each subpopulation." (NMFS, 2017).

Much of the Recovery Plan refers to data gaps and the need for gap analyses. Completing this necessary additional data-gathering delays answering key questions and implementing recovery actions. My research responds directly to this step of the recovery process—I address one data gap by laying the framework for a model of abundance in the Columbia River and suggesting the next research priorities for modeling how eulachon respond to ocean conditions.

A theme of the Recovery Plan is that the study of ocean conditions and eulachon abundance intersects in an area of limited data for eulachon with immense pressures on habitat and from climate change (NMFS, 2017). However, there is much data for the ecosystem that they occupy: the California Current upwelling zone (Peterson et al. 2014) and about the environmental conditions in the ecosystem that drive population dynamics of salmon (Burke et al. 2013). Salmon and eulachon are geographically and temporally linked, so it stands to reason that ocean conditions that drive salmon abundance and survival may similarly drive eulachon abundance and survival.

I suggest that ecological knowledge about eulachon as a forage fish as well as indicators developed for salmon can inform the processes in the ocean that drive eulachon abundance in a data-limited field. My research is predicated on the theory that patterns of eulachon abundance are not random but respond to physical, chemical, and/or biological factors such as temperature, prey abundance, or the Pacific Decadal Oscillation (as suggested in the Recovery Plan). I take this one step further, suggesting that eulachon abundance can be predicted by ocean conditions. This theory is supported by data for a similar genus, *Oncorhynchus*, or the Pacific salmon (NOAA, 2019). Because eulachon have a similar life history to salmon but are lower in the food web, and they have evolved under the same flow regimes in tributaries, I suggest that eulachon respond to some but not all of physical, biological, and oceanographic ecosystem indicators that Pacific salmon do. I suggest that some ocean conditions are correlated with higher abundance in a major eulachon spawning river (the Columbia River), and some conditions correlate with lower abundance.

The question of how ocean conditions drive eulachon abundance patterns in the Columbia River motivates my study. In exploring the relationship between ocean conditions and eulachon, I began with predetermined ocean ecosystem indicators for salmon. NOAA's Ocean Ecosystem Indicators (NOAA, 2019) were developed for Pacific salmon to understand how ocean survival is related to marine environmental conditions. Because these data are already being collected and are well understood in their relationship with salmon, using this ocean ecosystem indicators dataset was an efficient way to look at another piece of the food web—eulachon.

Complex ecological processes as well as anthropogenic effects determine the trajectory of species recovery. While the threats to eulachon in freshwater systems are less ambiguous than climate change impacts on ocean conditions, the population drivers during their ocean residency are not (Gustafson et al., 2010). There is no quantitative model of eulachon abundance that incorporates ocean conditions, and which could be used to inform fishery and recovery decisions. This analysis aimed to take a first look at some of these complex physical, chemical, and biological processes in the ocean that are already modeled for salmon and determine if the framework can be applied to eulachon. Specifically, I addressed the following questions. First, are the ocean ecosystem indicators that are used for predicting salmon abundance also correlated with eulachon abundance in the Columbia River? Second, which indicators drive this relationship and when?

Methods

Data

I used two datasets from the west coast of the U.S.: eulachon abundance data in the Columbia River and ocean ecosystem indicators from NOAA's Estuarine and Ocean Ecology program. The Columbia River and its tributaries make up one of the major spawning populations of the Southern DPS (NMFS, 2017). Spawning stock biomass estimates are used to estimate the number of adult fish returning to spawn. While there are other sources of eulachon abundance data, such as abundance data from the Fraser River from 1997 to present, and bycatch data from the pink shrimp fishery, I chose the Columbia River dataset because of the potential importance of the river to species recovery. Of rivers in the U.S., the Columbia River and its tributaries produce the majority of eulachon (Gustafson et al. 2010), and it is also more proximal to the Newport Hydrographic Line, one of the sources of ocean ecosystem indicators data, than other subpopulations considered by NOAA. Because my research question is directed at whether eulachon abundance is correlated with the same ocean ecosystem indicators as salmon, I used the same indicators (with few exceptions) as described by NOAA in the salmon forecasting model (NOAA, 2019).

The ocean ecosystem indicators data span 1998 to 2017. There are 16 indicators, summarized in Table 1. These data include biweekly data collection efforts off the Newport Line in Oregon, used to understand how marine environmental conditions affect juvenile and adult salmon survival. There are 3 Physical indicators, 4 Temperature indicators, 1 Chemical indicator, and 8 Biological indicators. The indicator data have different units, all of which are described in NOAA's analysis (NOAA, 2019). The dataset was downloaded from NOAA's Ocean Ecosystem Indicators website for use in this analysis; however, I modified the biological transition indicator to make it suitable for statistical modeling, as described in Appendix A. I also trimmed the dataset to start in year 2000 to match the eulachon abundance data.

The eulachon abundance data in the Columbia River span 2000 to 2017 and range from an estimated mean of 783,400 fish in 2005 to 185,965,200 fish in 2014, with 2017, the most recent year of data, seeing at estimate of 18,307,100 fish. From 2010 to 2017, abundance was determined from spawning stock biomass estimates. The same dataset includes ichthyoplankton abundance and eulachon sex ratios and fecundity estimates (Mallette, 2014). In the 2000 to 2010 dataset, eulachon abundance is estimated based on back-calculations using historic larval density data (NMFS, 2017). I acquired these data from NOAA (2017). Data from 2004 were missing and imputations of the missing data (see Appendix A) were not significantly predictive so 2004 was left out of the model. I also log-transformed the abundance data so that it resembled a normal distribution, necessary for use in linear modeling. The variability and (what appears to be) stochasticity of abundance in the Columbia River emphasize the use of multivariate techniques to understanding their life histories in the ocean.

Statistical modeling

One option to understand the relationship between abundance and indicators would be to perform linear regression of the individual indicators by eulachon abundance. However, I was curious about the overall effect of ocean conditions on abundance, not a single indicator's effect. Since the indicators are related and show multiple collinearity, it was appropriate to choose a technique that addresses the interaction between indicators. One such technique is PCR, which first determines a structure from complex datasets, then regresses that structure against a response variable (Legendre and Legendre, 1998). The complex dataset in this case was ocean ecosystem indicators; the PCR thus quantified the relationship between ocean ecosystem indicators and eulachon abundance. In PCR, the first step is to conduct a principal components analysis (PCA), which summarizes the variation in different indicators into fewer axes. Next, the PCR quantifies the relationship between the new axes and the other dataset of interest, in this case eulachon abundance. Not only is PCR a statistically appropriate method for addressing the relationship between eulachon and ocean conditions, it also allowed for comparison to modeling that has been performed for salmon (e.g., Burke et al., 2013).

I ran the PCA on a correlation matrix of the ocean ecosystem indicators data. I used a correlation matrix because the indicators have different units and scales. The correlation matrix treats the different indicators the same, with a standardized mean of zero and standard deviation of 1. The PCA produces principal components (PCs) which are uncorrelated new axes, where the first axis summarizes the dominant trends in variation and the second axis accounts for residual variance not accounted for by the first axis, and so on for additional axes. Having uncorrelated PCs resolved the issue of multicollinearity in ocean ecosystem indicators. For example, one would expect indicators such as PDO and temperature at multiple depths to be correlated because the PDO is a measure of sea surface temperature anomalies. Two biological indicators are also easily identifiable culprits for collinearity: the composition of Northern copepod species can be expected to inversely correlate with the composition of Southern copepod species because they are transported by opposing ocean currents. To determine how many PCs summarize significant amounts of variation from the original data, I tested whether the amount of variation described by each PC is higher than expected by chance. I used a broken-stick model (see Appendix A) to select the first two PCs as significant for further modeling.

The next step in PCR is relating the scores for the chosen principal components to eulachon abundance. To determine how eulachon abundance is related to ocean conditions, I performed a linear regression analysis of eulachon abundance against the PC scores for the first and second PCs. However, because I was interested in eulachon response to ocean conditions for the multiple years they reside in the ocean, I added time lags to the indicator data for a total of five analyses, testing the following hypotheses:

- Ocean conditions in the year that eulachon spawn (time lag 0 years) are correlated with abundance for that year
- Ocean conditions one year prior to spawning (time lag 1 year) are correlated with abundance
- Ocean conditions two years prior to spawning (time lag 2 years) are correlated with abundance
- Ocean conditions 3 years prior to spawning (time lag 3 years) are correlated with abundance
- Ocean conditions 4 years prior to spawning (time lag 4 years) are correlated with abundance

To determine how well the linear models fit the eulachon abundance data, I calculated the R^2 value and p-value for each regression. I used AICc (appropriate for small sample sizes) to evaluate model fit for models with significant results, which is further described in Appendix A. All analysis was conducted in R (R Core Team, 2019).

Results

The first section of the results addresses the first piece of the analysis—summarizing variation in ocean ecosystem indicators using PCA and identifying indicators that are representative of variation in the ocean conditions included in this analysis. The second section addresses the question of how this variation relates to eulachon abundance at varying time lags.

Summarizing variation in the ocean ecosystem indicators

A biplot of the PCA (Figure 2), shows the relationship between indicators, years, and the first two principal components (represented on the axes). Many of the same indicators that load heavily on the first PC have a similar direction and magnitude on the biplot, showing how these are correlated. For example, when the PDO is in a positive phase, the ONI is also positive, temperatures are relatively warm, and the copepod community is dominated by Southern species (Peterson et al., 2014), such as in 2015 and 2016 in the upper right corner of the biplot. In the opposite corner, years with cooler ocean conditions are present (2011 to 2013).

The PCA output describes how much variation each PC explains, as well as how much each indicator weighs on that axis. Because there was a high degree of multicollinearity in the indicator data, the variation in indicators was efficiently reduced to fewer axes. The first principal component (PC1) explains 51.8% of the variability in ocean ecosystem indicators, and PC2 explains 16.2% (Figure 3). Together, the first two PCs explain 68% of the variation in ocean ecosystem indicators. Within the first principal component, not all indicators contributed equally to the variation. Figures 4 and 5 show the contributions of the 16 indicators to the PC1 and PC2 scores. For PC1, the eight indicators on the left side of the graph had higher contributions than would be expected if all indicators loaded equally on the PC score—these indicators include bottom-up biological indicators like copepod biomass and abundance, as well as large-scale ocean conditions such as PDO and ONI. Indicators that contributed less to PC1 include site specific measurements such as temperature at the Newport line, and variables related specifically to salmon abundance (juvenile catches for Coho and Chinook). Table 1 shows the eigenvectors for each variable on PC1 and PC2. Eigenvectors are a measure of the magnitude and direction of the coefficients for each indicator. These same eight indicators that had higher contributions to PC1 and PC2 are color-coded in the table because they have the highest magnitude eigenvectors, representing their positive or negative loading on the PCs. The status of the PDO and ONI as well as the richness and composition of copepods and biomass of ichthyoplankton all have relatively high magnitude eigenvectors. These trends indicate more of the variation in the ocean ecosystem indicators is explained by large-scale patterns in ocean conditions and by species that respond directly to upwelling conditions (copepods), while less of the variation is explained by patterns at higher trophic levels (salmonids) or smaller geographic scales (such as sea surface temperature at the Newport Line). The next part of the analysis links eulachon abundance to the annual variation in conditions described by the PCA.

Relating ocean ecosystem indicators to eulachon abundance

The linear regression of eulachon abundance by PC1 and PC2 tells us whether the 68% of variation explained by PCs 1 and 2 is correlated with eulachon abundance, and the strength and statistical significance of the relationship. Table 2 shows the results from the PCR for each time lag scenario. In a linear regression that compares the compressed ocean ecosystem indicators variables (PC1 and PC2) against eulachon abundance with no time lag (i.e. the indicators are measured in the same year the spawning stock biomass is measured), there was no significant relationship. When the ocean ecosystem indicators are lagged by one year, there was a positive correlation with eulachon abundance ($p<0.05$, $R2=0.22$ on PC1). In this model run, variation in ocean conditions explained by PC1 accounts for some of the variation in eulachon abundance. Using a time lag of two years saw much stronger relationships. PC1 was again positively correlated with eulachon abundance ($p<0.001$) with a stronger relationship ($R2=0.55$). Using a time lag of three years saw similar results to the two-year time lag ($p<0.002$ and $R2=0.54$). The four-year time lag did not have significant results. The F-statistic, which tests the overall significance, was less than 0.05 for the two- and three-year time lags indicating these models outperform the other models.

Variation in ocean conditions explained by PC1 accounts for around half of the variation in eulachon abundance when lagged by two years and by three years. A positive residual median shows that the model underestimates eulachon abundance, on average. For all models, there were no significant results when modeling PC2. Indicators that load heaviest on PC2 include measures of temperature and salinity, and these are not correlated with eulachon abundance. However, this does not negate the influence of indicators that load heavily on PC2 on the survival of eulachon in the ocean –it is possible that these indicators are not significant in the model because they are already largely described by those that load heavily on PC1 (e.g. the PDO is also a measure of temperature). This means that ocean ecosystem indicators that contribute heavily to PC1 are most important in predicting eulachon abundance in any year. A comparison of AICc values for the two and three-year models evaluating just PC1 vs PC1 and PC2 also supports this conclusion (see Appendix A); both the two- and three-year models are more parsimonious using just PC1. In summary, the variation in ocean conditions that is represented by bottom-up and large-scale indicators is correlated with eulachon spawning abundance two years and three years later.

That the two- and three-year time lag models of spawner abundance are most parsimonious suggests that eulachon life histories in the Columbia River basin are dominated by fish that spend two to three years in the ocean; and data on the age structure of spawning populations provides empirical evidence for this variability in age structure in the Columbia River spawning population (Gustafson et al., 2010). Based on these modeling results, I illustrate a highly simplified version of the dominant life histories in the Columbia River (Figure 6), restricted to two- and three-year fish that spawn in late winter (the peak of a wide spawning period). Although it omits much of the variability in spawn timing, this figure nonetheless illustrates the complexity of understanding age class structure and eulachon in the ocean. In any year, the dominant age classes in the ocean could represent fish from five different life history or spawning strategies just from the Columbia River.

Discussion

The ocean ecosystem indicators dataset used for salmon is positively correlated with eulachon abundance in the Columbia River and should be used for future modeling. Eulachon respond to some of the same ocean ecosystem indicators as salmon. When eulachon are in the ocean when it is a cooler phase, such as in 2011 to 2013, returns two to three years later are higher than when the ocean is in a warmer phase. Salmon respond positively to the same ocean conditions (Burke et al., 2013). This is not a surprise due to their ocean residence, similar life histories, and trophic interactions, but nonetheless has not been quantitatively described in the literature. Understanding the relationship to indicators developed for salmon is the first step in building a customized model for accurately and precisely predicting eulachon abundance in the Columbia River or throughout the Southern DPS. Understanding drivers of abundance will also help inform eulachon recovery planning and prioritize recovery actions, in addition to inform their ecological patterns. This work also provides a methodological contribution to the field of marine ecology in its transferal of data for salmon to a related species. And it occurs at the intersection of science and policy, where choices that are made about sensitive species reflect the priorities of people who have power to make decisions as well as historical imbalances in data collection.

I found that ocean conditions in the three years prior to return are predictive of eulachon abundance; data from two years and three prior have a stronger and more significant relationship than one year prior which suggests dominant life history strategies in the Columbia River. One of the interesting results from this analysis is that indicators related to bottom-up processes and large-scale oceanic drivers such as the status of the PDO are important to eulachon abundance. However, other processes not included in this analysis likely also influence Columbia River eulachon spawning abundance, such as top-down processes like predation in the estuary and bottom-up processes like stream flow regime. Marine mammals and piscivorous birds predate on adults returning to spawn, but data are lacking regarding the effects of these populations (NMFS, 2017). Freshwater and estuary conditions likely affect spawning, rearing and outmigration due to effects of temperature and water quality on eggs and juveniles, habitat quality, predation, and other processes, but little data documents these effects (NMFS, 2017). A life cycle model incorporating riverine and estuarine factors may explain some of the 45% of variance that is not accounted for in my model, which uses only ocean conditions.

Bottom-up biological indicators like copepod biomass and abundance, as well as major climatic/ocean indices (PDO, ONI) that load heavily on PC1 are likely key drivers of patterns in eulachon abundance in the Columbia River. These indicators efficiently represent the overall variation that occurs within the indicators dataset. Understanding how ocean conditions influence productivity can improve forecasts of how continuing climate change may affect eulachon. Climate change impacts on ocean conditions were classified as the most serious threat to the Southern DPS of eulachon by NOAA's Biological Review Team and were identified as a high research priority in the Recovery Plan (Gustafson et al., 2010; NMFS, 2017). The upwelling patterns that determine productivity along the west coast could change and shifts in zooplankton communities are likely, which could cause a decline in forage fish populations or distributional changes in eulachon compared to preferred prey (ISAB, 2007).

Suggestions for future research

Suggestions for future research are summarized below in two categories: 1) improvements to the "proof of concept" model described in this paper, and 2) additional modeling efforts for freshwater and estuary conditions to inform how environmental or biological factors influence eulachon a different life history stages.

1. Fine-tune salmon indicators for eulachon

   Using the ocean ecosystem indicators for salmon was a first attempt to identify whether a relationship exists between ocean conditions and eulachon and identify some of the indicators that appear to drive this relationship. The suggestions below describe how the indicator dataset could be handled, analyzed, or modified in a different way to improve the model.

   1.1. Combine ocean conditions from multiple years into model for eulachon abundance in a single year

This statistical model focused on the effects of a single year of ocean conditions on a single year of eulachon abundance. However, eulachon spend multiple years in the ocean and therefore affected by multiple years of ocean conditions. Future work should explore how multiple years of ocean conditions (the two to five years prior to return) can be combined to predict abundance. Averaging the conditions for these years and performing a linear model may not be appropriate due to violations of the autocorrelation assumption. More robust time series analysis could inform which years of ocean residency are most influential. Combined with a better understanding of age-class structure, this could inform why eulachon return in a specific year.

1.2 Incorporate age class structure into model

This model does not incorporate age class structure. In any year, eulachon on spawning grounds could represent fish that have spent two to five years in the ocean, with some exceptions (NMFS, 2017). Using results from otolith aging or alternative methods, the model could be improved by understanding the proportion of each brood year represented in that return year, and how this varies year to year in response to ocean conditions.

1.3 Make ocean ecosystem indicators specific to eulachon timing of ocean residency

Future research should trim the ocean conditions data to the time of year when eulachon are theorized to be present in the upwelling system. For example,

1.4 Determine if some indicators are not predictive in the model (reduce number of indicators)

Using the indicators developed for salmon has a disadvantage in that I assumed these indicators are appropriate to use for eulachon. For example, two indicators are specific to juvenile salmon (catches of Chinook and Coho juveniles in June), and these did not contribute heavily to PC1. Future analyses, such as a leave-one-out analysis, may inform which indicators for salmon can be removed to better model eulachon abundance.

1.5 Incorporate additional ocean data (add new physical/chemical indicators, if available)

The model includes many large-scale ocean condition indicators as well as more local indicators that were developed for the salmon models. The local indicators are less predictive than the large-scale indicators for eulachon. Future research could search for additional ocean data and incorporate new indicators, testing whether the model becomes more predictive with each iteration.

1.6 Incorporate top-down ecological processes specific to eulachon (add new biological indicators, if available)

Similar to the previous suggestion, there may be biological indicators in the ocean that were not included in the salmon model that may be informative for eulachon abundance. While the model incorporates bottom-up processes by including indexes of eulachon prey base, it does not incorporate many eulachon predators in the ocean. Future research could search for additional data related to top-down ecological processes and incorporate new indicators, testing whether model performance improves.

2 Model freshwater or estuary conditions that influence survival and abundance

This analysis focuses only on the ocean ecology of eulachon. To build a more complete understanding of how environmental conditions affect populations of eulachon, data from the freshwater and estuary systems should be incorporated. The suggestions below describe how non-ocean conditions could be studied to ultimately inform a life cycle model for eulachon.

2.3 Hydrograph for Columbia River (or tributaries with significant spawning populations)

Water management in the Columbia River basin affects the flows that eulachon experience during migration, spawning, and rearing. Future research could focus on how the hydrography of the Columbia River as well as major spawning tributaries (such as the Cowlitz River) affects processes such as larval growth, juvenile development and migration, and survival in the estuary environment.

2.4 Temperature data for Columbia River (or tributaries with significant spawning populations)

Water management and climate change affect temperatures that eulachon experience during migration, spawning, and rearing in the freshwater environment. Like the previous suggestion, future research could incorporate temperature data from the Columbia River or tributaries to models of eulachon survival in the freshwater system.

2.5 Plume conditions in the Columbia River estuary

The observation and modeling network SATURN reports physical and biological conditions in the lower Columbia River estuary (NMFS, 2017). Combined with sampling efforts, these data could be used to understand how eulachon respond to changes in the plume environment and how plume conditions affect growth and survival, as previously suggested in the Eulachon Science to Policy Forum (Anchor QEA, 2015).

Improving the model and building a life cycle model could assist recovery planning by providing information on the critical physical, chemical, and biological features that eulachon need to survive at different life stages. It would help researchers decide which indicators need more data collection and which life stages or habitat conditions need the most protection.

Contributions to the field

Harnessing and analyzing data is one small intervention to eulachon recovery by addressing a key knowledge gap, but multivariate statistics is not the solution to eulachon recovery. Addressing their major threats is. Understanding how variability in ocean conditions will affect eulachon recruitment and survival is important to the recovery process, but just understanding these relationships will not change the ocean conditions for the better or increase eulachon populations. Eulachon are a critical component of freshwater, estuary, marine, and human food webs, so interventions are possible in other aspects of their ecology. A model that incorporates each component of eulachon ecology will provide the greatest potential for identifying actionable recovery steps.

Though, species recovery processes rarely advance in a linear direction between data collection, analysis, and management recommendations. Especially for eulachon, being understudied exacerbates the major threats like climate change and habitat alteration. When the priority actions listed in a recovery plan are identifying data gaps and collecting more data, there is a time lag to making policy changes or implementing recovery-oriented projects that could be identified by this additional knowledge. And addressing one gap highlights other areas where more work is needed. This analysis has taken available data and learned something new about eulachon ecology, and the results highlight that more kinds of data, and additional years of data, from ocean and freshwater systems are needed. As suggested by NOAA in the Recovery Plan, a marine abundance estimate would significantly improve understanding of eulachon ecology in the ocean and where bottlenecks to their survival are. NOAA (2012) examined the potential to develop a marine abundance estimate and found that the available marine data (mostly from trawl surveys) showed discrepancies between marine catches and freshwater estimates, leading the authors to consider that marine indices may represent catchability rather than availability. For this reason, they considered freshwater abundance metrics, where eulachon are counted in a defined area, to more likely represent the status of eulachon. But freshwater indices have their own drawbacks; an important one to understanding how eulachon will respond to major threats like climate change is: where does marine mortality occur and why? This model does not answer this key question but suggests that prey composition and abundance, which is influenced by upwelling and oceanic patterns like the PDO and ONI, may be a key area of further study.

This work implies that a multi-species modeling framework including these indicators, salmon, eulachon, and other species for which abundance data are available may be helpful for understanding how individual species and the interactions between them respond to changes in ocean conditions. Multivariate statistics were a supremely economical and useful tool for the question at hand, using only available data – meaning that understudied species can benefit from work that is completed for more well studied species in cases where the two have much in common. This economical modeling effort has proven that ocean ecosystem indicators that were developed for salmon are a reasonable starting point for modeling eulachon abundance in the Columbia River and in the Southern DPS. I do not advocate for a blanket transferal of modeling work from one species to another; however, the ocean ecosystem indicators developed for salmon represent overall conditions that salmon experience in the California Current and are suitable to more than just salmon. Peterson et al. (2014) previously demonstrated this with sablefish and sardines. They described how zooplankton biodiversity is determined by ocean currents and large-scale factors like the PDO and ONI, affecting species such as salmon at higher trophic levels. Their findings and my own bolster the body of literature suggesting that NOAA's ocean ecosystem indicators have utility for understanding and modeling the California Current ecosystem as a whole. And my work further suggests that for fishes in the Columbia River and especially eulachon, incorporating more top-down drivers such as estimates of avian and marine mammal predation, as well as environmental conditions in the freshwater and estuary environments may also be important to addressing their major threats. My work is further evidence that modeling frameworks that incorporate the complex ecological interactions which determine recruitment and survival at different trophic levels may help recover populations that depend on marine and freshwater ecosystems.

Acknowledgements

I would like to thank my committee: Dr. Cleo Woelfle-Erskine and Robert Anderson for their support, encouragement, and comments regarding this work. Robert has provided key input and connections to NOAA staff plus a shared curiosity for eulachon. I especially thank Cleo who has advised and mentored me in my time at SMEA. Under Cleo's wing I have grown as a scientist, writer, and thinker. I am grateful for his wisdom, clarity, and kindness.

Thank you to Dr. Rick Gustafson at NOAA for providing vital brainstorming. I am grateful for his expertise and prior work on eulachon ecology and modeling. I hope this work is a small contribution to the recovery process.

Thank you to the FRESH (Feminist Research in Eco-Social-Hydro) Water Relations Lab for being my intellectual home. What a bright home to share and build with you all—thank you for your edits and ideas, your commiseration and enthusiasm. Thank you to my SMEA cohort and classmates. You have talked me through analyses, drafts, intersections, and our lives as scientists and humans. Sharing this space (intellectual, physical, and lately, virtual) with you has been a joy and I hope it is just the beginning of our time together.

I am thankful for the support and encouragement of many faculty at SMEA and at UW. Dr. Ryan Kelly and Dr. Julian Olden both provided friendly guidance with R and statistical modeling. In addition, SMEA's array of multidisciplinary classes have let me explore aspects of marine affairs entirely unrelated to eulachon, for which I am grateful.

Thank you to Anchor QEA for the tuition assistance and flexibility while I have been in school. Thanks especially to the natural resources and fisheries teams who make coming to work every day such a joy. A special thank you to Dr. John Ferguson for introducing me to eulachon, mentoring me throughout my early career, and for inspiring this analysis. And thank you to John Small for your mentorship and support. I would not be here without you both.

I appreciate all of my friends who have cheered me on both personally and scholastically while I've been in school. Thank you for asking about my work, checking on my knee, and being a source of light at the end of the tunnel.

Thank you deeply to my family for their unconditional support. Thank you particularly to my mom, Elaine, and dad, Bob. Mom, your unrelenting support of my education and physical and mental health has given me the tools to be resilient in this world. Dad, I have wanted to learn more about fish since I was a little girl all because of you—thank you for sharing this joy with me and supporting me every step of the way. Thank you also to my brothers, Jack and Michael, friend Ariel, and dogs, Quincy and Mak, for having distractions and kindnesses at the ready. Thank you to my family for providing a home at Whidbey where I first fell in love with the ocean and became an ecologist.

Finally, thank you eulachon! You reignited my curiosity and drive for studying marine ecology and understanding the relationships between humans and the natural world. You are the most interesting puzzle I have ever encountered.

References

Anchor QEA, 2015. Summary of the Northwest Power and Conservation Council's Eulachon State of the Science and Science to Policy Forum. Prepare for: Northwest Power and Conservation Council, Portland, OR. October, 2015.

Broduer, R.D., TD. Auth, and A.J. Phillips. 2019. Major shifts in pelagic micronekton and microzooplankton community structure in an upwelling ecosystem related to an unprecedented marine heatwave. *Frontiers in Marine Science,* 6:212.

Burke, B.J., W.T. Peterson, B.R. Beckman, C. Morgan, E.A. Daly, and M. Litz. 2013. Multivariate models of adult Pacific salmon returns. *PLoS ONE*. 2013:8(1).

Federal Register, 75:13012. Edangered and Threatened Wildlife and Plants: Threatened Status for Southern Distinct Population Segment of Eulachon. Department of Commerce, NOAA.

Gustafson, R.G., M.J. Ford, D. Teel, and J.S. Drake. 2010. Status review of eulachon (*Thaleichthys pacificus*) in Washington, Oregon, and California. NOAA Technical Memorandum NMFS-NWFSC-105.

ISAB (Independent Scientific Advisory Board). 2007. Climate change impacts on Columbia River basin fish and wildlife. ISAB, Report 2007-2, Portland, Oregon.

Legendre, P. and L. Legendre. 1998. Numerical ecology, Elsevier.

Leland, D., and Mitchell, D. 2001. Sharing resources on the North Pacific coast of North America: the case of the eulachon fishery. *Anthropologica*, *43*(1), 19–35.

Litz, M., R. Emmett, P. Bentley, A. Claiborne, and C. Barcelo. Biotic and abiotic factors influencing forage fish and pelagic nekton community in the Columbia River plume (USA) throughout the upwelling season 1999-2009. ICES Journal of Marine Science, 71(1), 5-18.

Mallette, C. 2014. Studies of Eulachon Smelt in Oregon and Washington. Project Completion Report. Prepared by Oregon Department of Fish and Wildlife and Washington Department of Fish and Wildlife. Prepared for NOAA. Grant Number NA10NMF4720038. September, 2014.

NMFS. 2016. 2016 5-Year Review: Summary and Evaluation of Eulachon. National Marine Fisheries Service West Coast Region, Portland, OR.

NMFS. 2017. Recovery Plan for the Southern Distinct Population Segment of Eulachon (*Thaleichthys pacificus)*. 132. Prepared by: National Marine Fisheries Service, West Coast Region. Protected Resources Division, Portland, OR.

NOAA. 2019. Ocean ecosystem indicators of salmon marine survival in the Northern California Current. Retrieved from: https://www.nwfsc.noaa.gov/research/divisions/fe/estuarine/oeip/index.cfm.

Peterson, W. T., J.L. Fisher, J.O. Peterson, C.A. Morgan, B.J. Burke and K.L. Fresh. 2014. Ecosystem Indicators of Ocean Conditions Inform Fisheries Management in the California Current. *Applied Fisheries Oceanography*: 27(4).

R Core Team. 2019. R: A Language and Environment for Statistical Computing. Vienna, Austria.

Tables

Table 1. Ocean ecosystem indicators (adapted from Peterson et al., 2014) and PC1 and PC2 eigenvectors. The 8 strongest weighting indicators for PC1 and PC2 are color coded, green for positive loading and red for negative.

Category	Indicator	PC 1 eigenvector	PC 2 eigenvector
Climate	PDO (Pacific Decadal Oscillation; December-March)	0.287994	-0.1111126
	PDO (May-September)	0.2973513	0.0600706
	ONI (Oceanic Niño Index; Average, January-June)	0.2980782	-0.095399
Temperature	46050 SST (degrees C, May-September)	0.2243998	0.2661159
	Upper 20m (degrees C, November-March)	0.3024022	-0.2046106
	Upper 20m (degrees C, May-September)	0.1485832	0.4749114
	Deep temperature (degrees C, May-September)	0.2188504	0.3302323
Chemical	Deep salinity (May-September)	-0.0781131	-0.4572665
Biological	Copepod richness anomaly (no. species, May-September)	0.3166222	-0.1694638
	Northern composition copepod biomass anomaly (mg C per cubic meter, May-September)	-0.2873284	0.1859103
	Southern composition copepod biomass anomaly (mg C per cubic meter, May-September)	0.3308438	-0.036495
	Biological transition occurred (Y/N)	0.2400455	-0.2455204
	Ichthyoplankton biomass (mg C per 1000 cubic meter, January-March)	0.0165891	-0.4121462
	Ichthyoplankton biomass index (PCO axis 1 scores, January-March)	0.3041277	0.0293726
	Chinook salmon juvenile catches (number per km, June)	-0.2370577	-0.1137909
	Coho salmon juvenile catches (number per km; June)	-0.151864	-0.1000309

Table 2. Statistical results for PCR of eulachon abundance and ocean ecosystem indicators

Analysis	**Residual Median**	**P value**		**F statistic (p-value)**	**Adjusted R-squared**
		PC1	PC2		
Time lag 0 years	0.3385	0.65	0.637	0.2039 (0.8574)	-0.1105
Time lag 1 year	0.0894	0.0271*	0.6437	3.168 (0.0876)	0.2243
Time lag 2 years	0.06215	**0.00089***	0.5725	9.673 **(0.00396*)**	0.5534
Time lag 3 years	0.00241	**0.00164***	0.37947	8.634 **(0.005565*)**	0.5401
Time lag 4 years	-0.003705	0.154	0.190	3.724 (0.06185)	0.3122

* denotes p value <0.05

Table 3. Results of model fit comparison using AICc

Analysis	**PCs used in model**	**AICc**
Time lag 2 years	PC1 and PC2	34.09404
Time lag 2 years	PC1	30.69098
Time lag 3 years	PC1 and PC2	33.17583
Time lag 3 years	PC1	29.95086

Figures

Figure 1. Map of major eulachon spawning populations with Southern DPS indicated in brackets, annotated from NMFS 2017

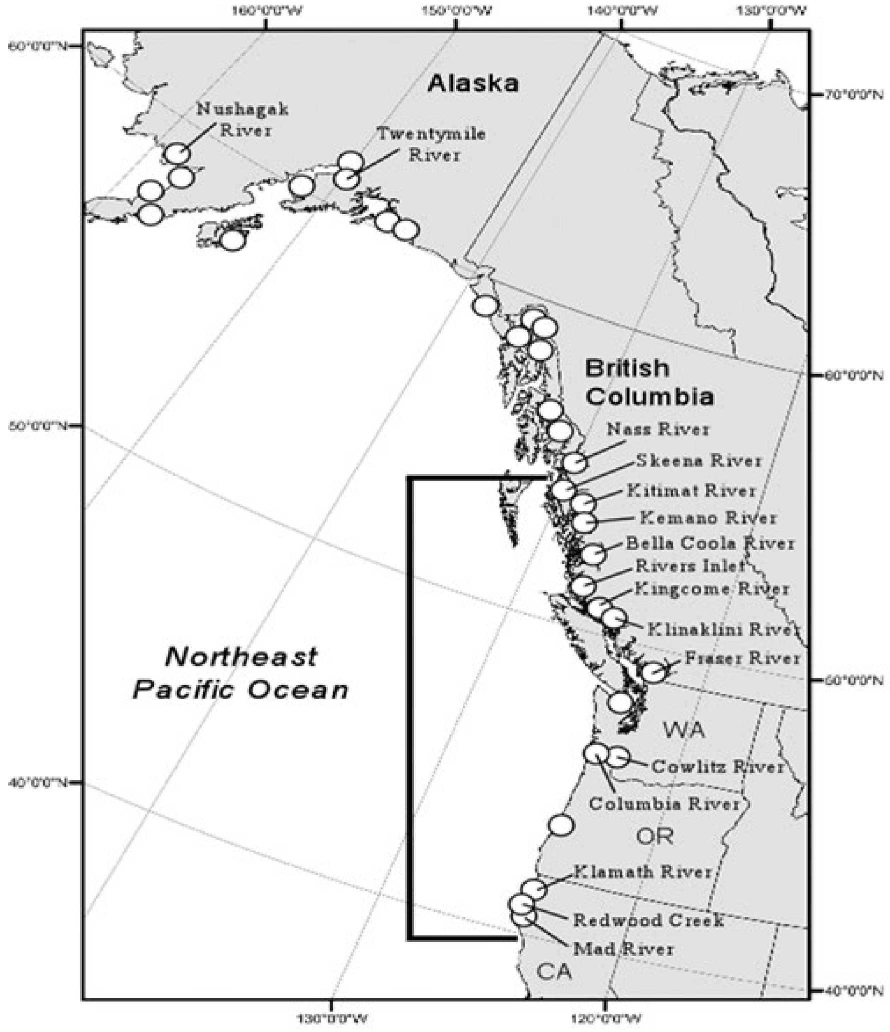


Figure 2. PCA biplot showing eigenvectors for PC1 and PC2 related to years of occurrence, with key indicators labeled in black

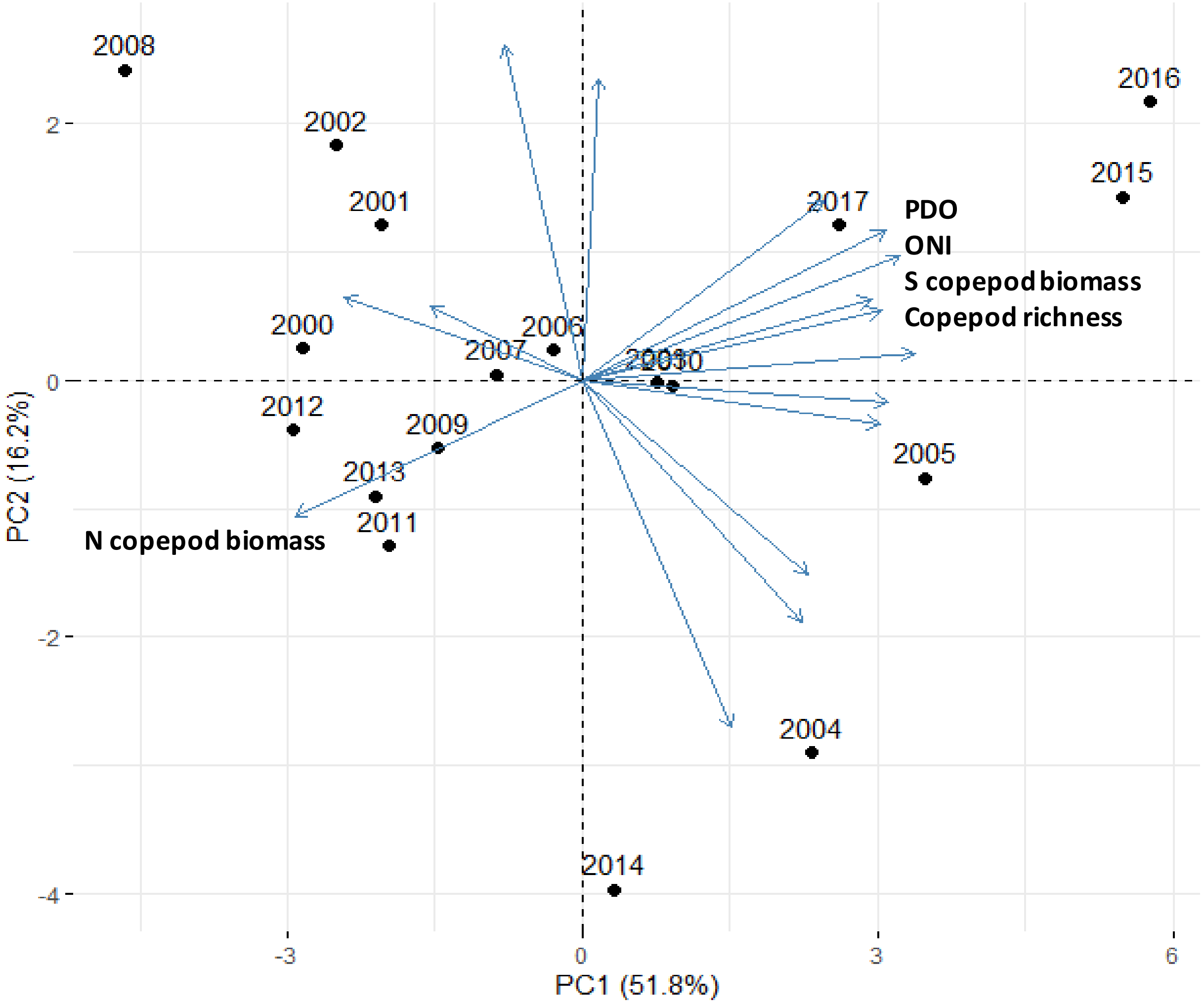


Figure 3. Percentage of variance explained by each PC

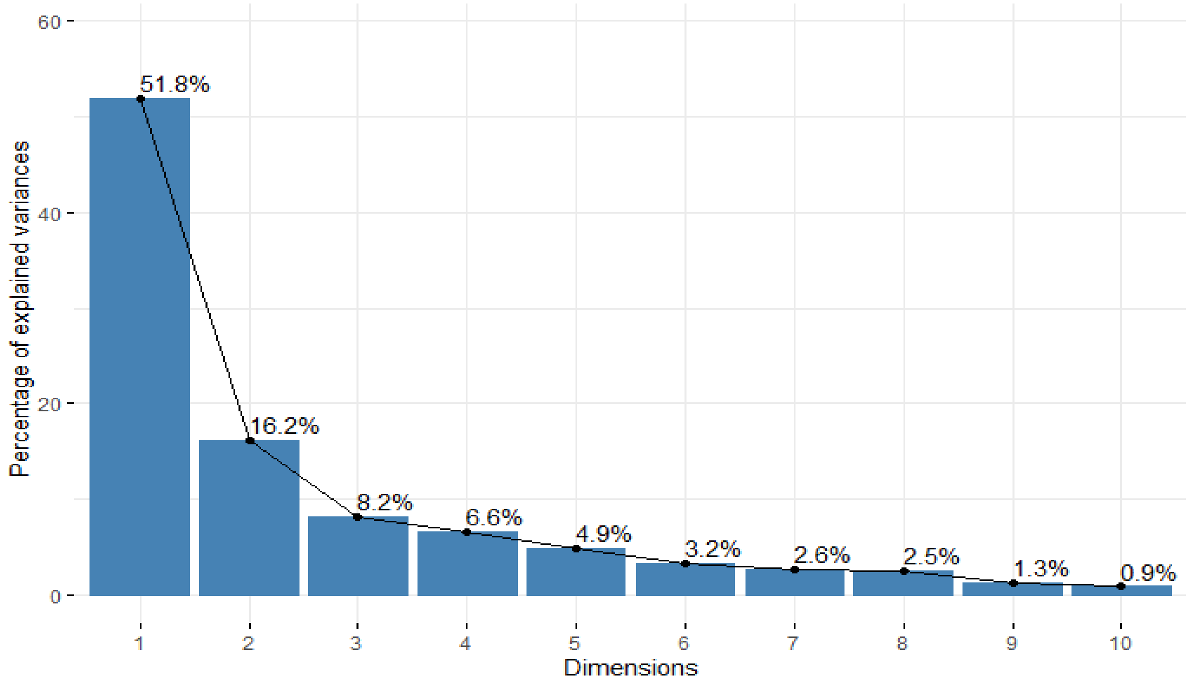


Figure 4. Contribution of each ocean ecosystem indicator to PC1. Red dotted line represents what contribution would be if all indicators had equal contribution

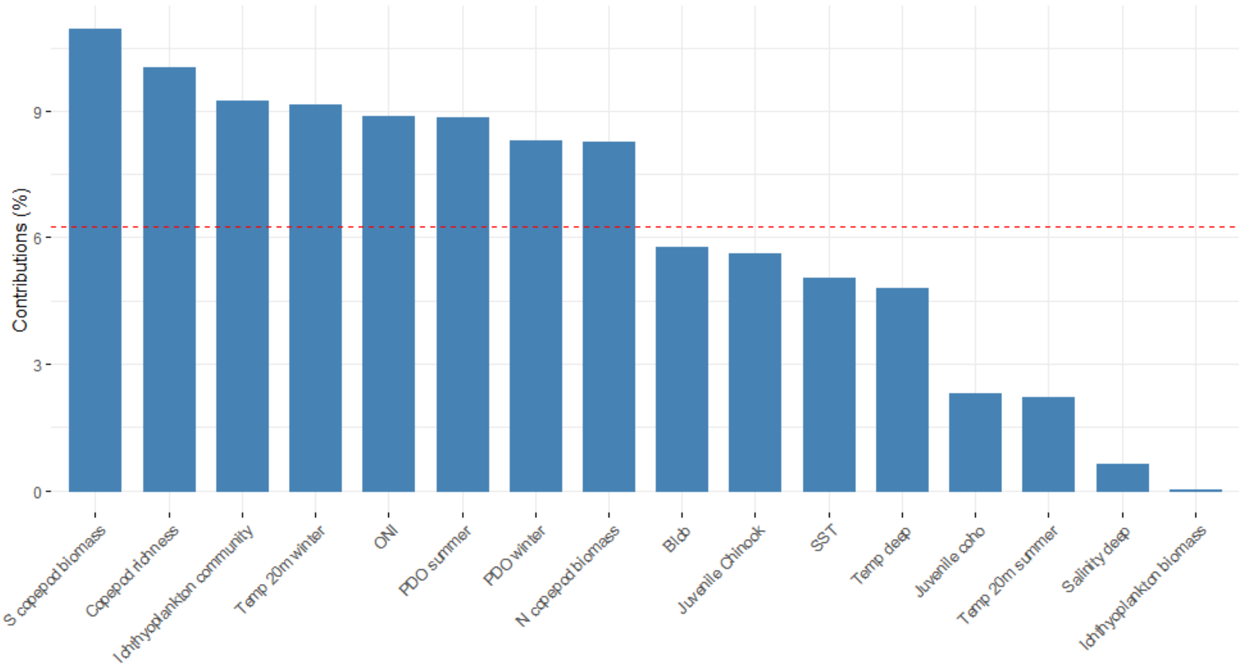


Figure 5. Contribution of each ocean ecosystem indicator to PC2. Red dotted line represents what contribution would be if all indicators had equal contribution

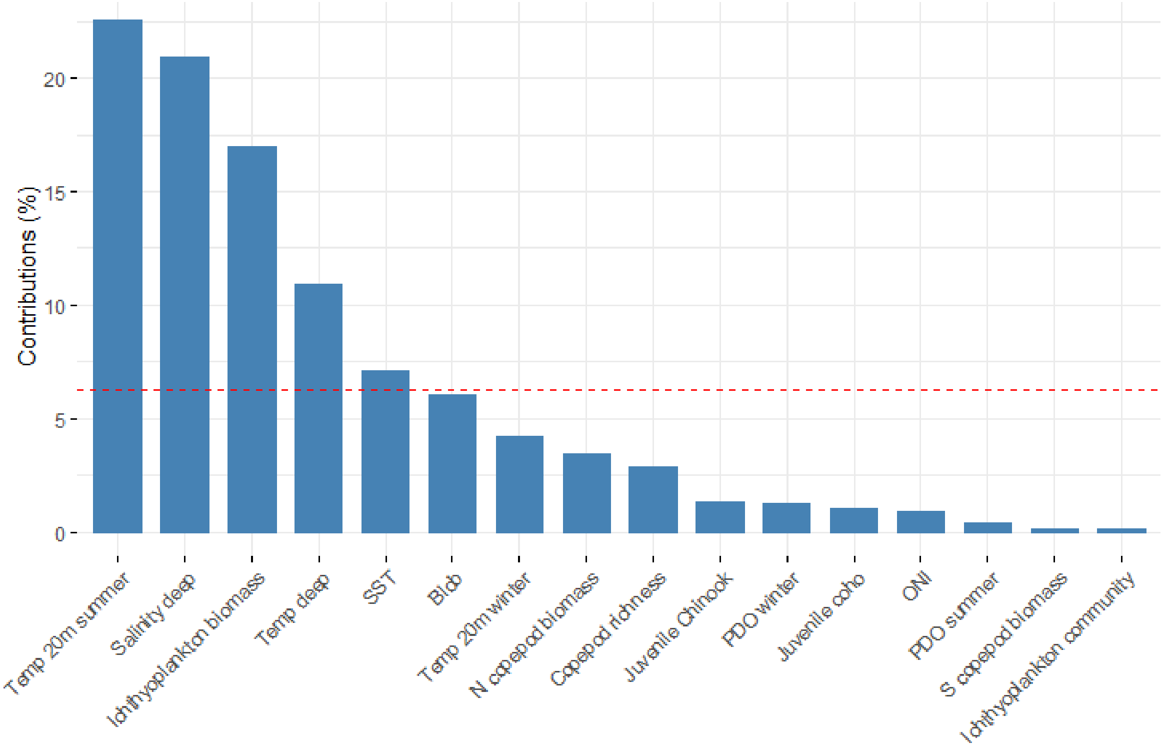
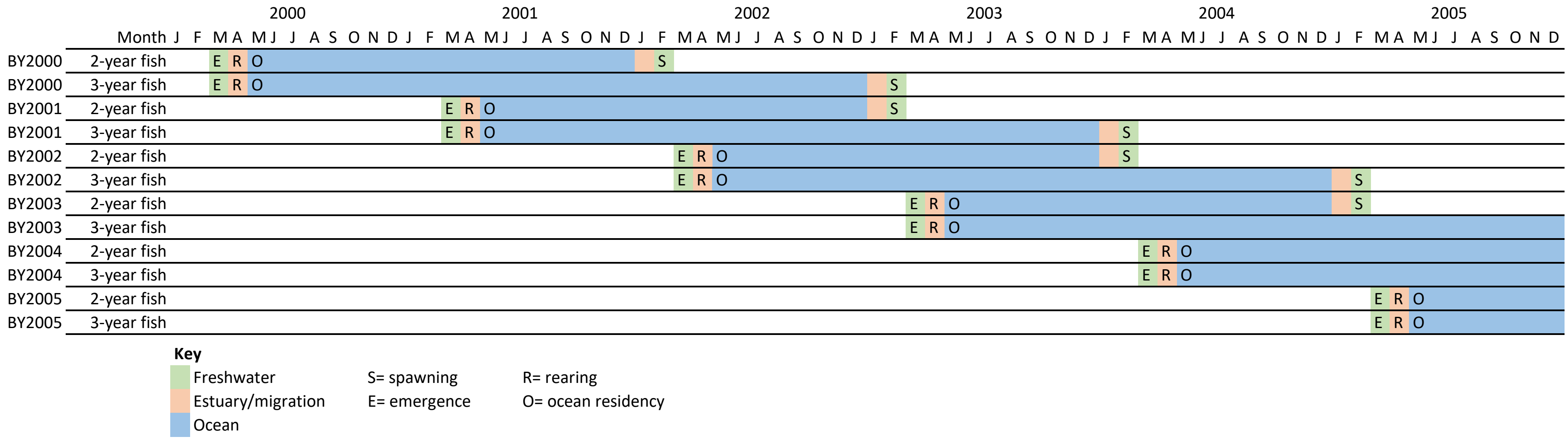


Figure 6. Proposed dominant life history of eulachon in the Columbia River, simplified example with brood years starting in 2000



Detailed life history timing (summarized from Gustafson et al. 2010)

S Spawning in the Columbia River occurs from late December to late April; February is the estimated peak of abundance

E Eggs incubate for approximately three to five weeks

R Larvae may spend weeks to months rearing in freshwater and estuary environments, denoted here as the month of April, but could feasibly occur from January to July

O Ocean residency could begin anytime from around February to July; denoted here as March

Appendix A: Analytical Supplement

There were numerous pieces of statistical modeling that are not suitable to include in a summary of this research. These items include details regarding the data sources, data screening procedures, selection of principle components for use in the regression model, and model fitting. These are described in the following sections. All analyses were conducted using R and the packages I used are cited at the end of this section.

Source Data

The main document includes information about the primary two sources of data for this analysis, ocean ecosystem indicators from NOAA and estimates of eulachon abundance in the Columbia River. Additional sources of data that were used during data exploration and screening are also discussed below.

Table 1. Data used in analysis

Data Source	Units	Years Used	Source of Data
Ocean Ecosystem Indicators	Variable	2000-2017	NOAA 2019
Eulachon Abundance in Columbia River	Mean estimate of run size	2000-2003; 2005-2017	NMFS 2017 Table 2-3
Eulachon Biomass in Fraser River	Pounds	2000-2017	NMFS 2017
Commercial harvest data from the Columbia River	Pounds	2000-2010	NMFS 2017

Ocean Ecosystem Indicators

This dataset is discussed in the Methods section of the main text.

Eulachon abundance in Columbia River

This dataset is discussed in the Methods section of the main text. In addition to what is described there, it is important to note some that this dataset represents a mean estimate of the number of run size, and this estimate is gathered from two different methods depending on the year. From 2000 to 2010, estimates were back-calculated using historical larval density data (NMFS, 2017). From 2010 to 2017, estimates were calculated based on spawning stock biomass surveys as summarized in Mallette (2014). Details regarding both methods and how the calculations were conducted are included in the Recovery Plan.

Eulachon biomass in Fraser River

The Fraser River biomass dataset comes from data provided from the Department of Fisheries and Oceans Canada (DFO) that were included in the Recovery Plan. The dataset spans 1997 to 2017 and provides an estimated run size of eulachon in pounds. These data were used to test whether the missing 2004 estimate for eulachon abundance could be imputed using Fraser River data as described below. This test was predicated on the hypothesis that the Fraser River run and Columbia River run would be correlated enough to use one to predict the other when a year is missing. Because the test was unsuccessful, this dataset was not used further in the analysis.

Commercial harvest data from Columbia River

The commercial harvest dataset in the mainstem Columbia River comes from data provided from WDFW that were included in the Recovery Plan. I chose to use the mainstem commercial data because the tributary, sport, and tribal estimates were lacking many years of data. These data were used to test whether the missing 2004 estimate for eulachon abundance could be imputed using commercial harvest data as described below. This test was predicated on the hypothesis that commercial harvest data and estimates of eulachon abundance would be correlated enough to use one to predict the other when a year is missing. Because the test was unsuccessful, this dataset was not used further in the analysis.

Data Screening Procedures

An important part of multivariate analysis is data screening and exploration. Data screening for this project involved identifying missing data, standardizing and transforming data, and treating missing data. The below sections describe how I screened my datasets.

Cleaning and standardizing

An empirical cumulative distribution function of the eulachon abundance data indicated that eulachon abundance was skewed. I performed a log-10 transformation of the eulachon abundance data before using it for analysis.

The ocean ecosystem indicators data have varying units. Setting “scale” to “TRUE” in the “prcomp” command scales the original variables to unit variance, which provides a correlation matrix. The correlation matrix standardizes values to a mean of 0 and standard deviation of 1, so there was no need to transform the ocean ecosystem indicators data.

Missing data for 2004

Abundance estimates for 2004 were not available because larval density data in 2004 in the Columbia River were not available to perform the back-calculations. Because the 2004 data were missing completely at random, I attempted to impute that value using other related eulachon run data from 2004. NOAA reports mainstem and commercial catch landings from 2000 to 2017 in the Columbia River (for which some years’ data are also missing, but 2004 is available), as well as the run size in the Fraser River from 1995 to 2017. I used a predictive algorithm to use the commercial landings and Fraser River data to impute multiple potential values for the Columbia River abundance in 2004; however, neither the commercial harvest nor the Fraser River values were significantly predictive of Columbia River abundance (p=0.57 and p=0.55, respectively). I performed this using the "mice" function, which stands for Multivariate Imputations by Chained Equations, and used an imputation method suitable for any category of variable. Not being able to predict an artificial value for the 2004 abundance, I used "NA" and excluded the 2004 abundance from the final statistical models.

Accounting for The Blob

Examination of the ocean ecosystem indicators data yielded one issue with the variable "Biological transition (day of year)." The biological transition indicator is the day of the year that a northern (cold-water) copepod community appears at the Newport station. This is a useful indicator because northern copepod species are lipid-rich and favored by coho and Chinook salmon. This indicator ranges from the 64th to the 214th day of the year, except in 2015 and 2016, when the biological transition did not occur. NOAA reports the value for this indicator as 365 in these two years, but this is an artificial value which does not describe the ecological implications of the transition not occurring and thus is not useful for statistical modeling. In 2015 and 2016, an unprecedented marine heat wave occurred off the west coast of the U.S., causing disruption to typical upwelling patterns and shifts in community structure (Broduer et al., 2019). I replaced the biological transition indicator with a binary indicator of whether the heat wave was occurring in that year (0 for not occurring, and 1 for 2015 and 2016). This allowed me to retain some of the information regarding the disruption to the biological spring transition without skewing data by using an artificial value for 2015 and 2016. I did this in the raw data file so code for this update is not provided in this appendix. I renamed this indicator "Blob," which represents the lower resolution of the binary values from the initial data.

Selection of principle components

Each principle component has an associated eigenvalue that represents the variance it explains. When an eigenvalue is higher for one PC than another, it explains more of the variance in the dataset. The broken stick model is a null model that generates a random distribution of eigenvalues. Comparing the eigenvalues from each PC to the eigenvalues of the broken stick model determines whether they are higher or lower than would be expected by chance, thus identifying the PCs which are considered to explain a statistically significant portion of the variance in the ocean ecosystem indicators dataset. Using the broken stick model, I selected the first and second PCs for further modeling.

Model-fitting

Evaluating the results of the PCR allowed me to narrow down the most descriptive models to time lags of 2 and 3 years using both PC1 and PC2 as described in previous sections of this paper. However, R^2 values and p-values do not tell us how the models compare with respect to how well the model fits and how simple it is. I used AICc, which is suitable for small sample sizes, to compare model fits. The results of the AICc are included in Table 3 of the main text and again below. Note that the two models using just PC1 had similar AICc values compared to the two models using PC1 and PC2 despite the differences in time lag. This means that the 2- and 3-year models perform similarly well when using PC1. This also agrees with the PCR results, where I saw no significant results when regressing eulachon abundance against PC2. Overall, the AICc comparison shows that the 2- and 3-year lags using PC1 are the best fitting models given the available data. This emphasizes my recommendation to examine the indicators that load heavily on PC1 more closely and to continue incorporating age class structure and variable life histories of eulachon when possible.

Table 2. AICc comparisons of models using PC1 vs. PC1 and PC2

Analysis	**PCs used in model**	**AICc**
Time lag 2 years	PC1 and PC2	34.09404
Time lag 2 years	PC1	30.69098
Time lag 3 years	PC1 and PC2	33.17583
Time lag 3 years	PC1	29.95086

References to Packages and Software Used

Grosjean, P. and F. Ibanez. 2018. pastecs: Package for Analysis of Space-Time Ecological Series. R package version 1.3.21. https://CRAN.R-project.org/package=pastecs.

Kassambara, A. and F. Mundt. 2017. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.5. https://CRAN.R-project.org/package=factoextra.

Le, S., J. Josse, and F. Husson. 2008. FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software, 25(1), 1-18. 10.18637/jss.v025.i01.

Mazerolle, M. 2019. AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c). R package version 2.2-2. https://cran.r-project.org/package=AICcmodavg.

McGarigal, K. 2014. biostats.R source file.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, R., McGlinn, D., Minchin, P.R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H., Szoecs, E and Wagner, H. (2019). vegan: Community Ecology Package. R package version 2.5-6. https://CRAN.R-project.org/package=vegan.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

RStudio Team. 2019. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. http://www.rstudio.com/.

Van Buuren, S., and K. Groothuis-Oudshoorn. 2011. Mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3), 1-67. https://www.jstatsoft.org/v45/i03/.