NOAA Technical Memorandum NMFS-AFSC-485

Methods for Combining Ichthyoplankton Datasets for Predicting Annual Larval Fish Abundances in the Gulf of Alaska

L. N. Block, R. R. Hopcroft, and L. A. Rogers

May 2024

U.S. DEPARTMENT OF COMMERCE

National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Alaska Fisheries Science Center

This document should be cited as follows:

Block, L. N., Hopcroft, R. R., and Rogers, L. A. 2024. Methods for combining ichthyoplankton datasets for predicting annual larval fish abundances in the Gulf of Alaska. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-AFSC-485, 20 p.

This document is available online at:

Document available: https://repository.library.noaa.gov

NOAA FISHERIES

Methods for Combining Ichthyoplankton Datasets for Predicting Annual Larval Fish Abundances in the Gulf of Alaska

L. N. Block[1], R. R. Hopcroft[2], and L. A. Rogers[3]

[1]Pacific Biosciences Research Center
University of Hawai‘i at Mānoa
1993 East West Rd.
Honolulu, HI 96822

[2]College of Fisheries and Ocean Sciences
University of Alaska Fairbanks
Fairbanks, AK 99775-7220

[3]Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
7600 Sand Point Way NE
Seattle, WA 98115

U.S. DEPARTMENT OF COMMERCE

National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Alaska Fisheries Science Center

NOAA Technical Memorandum NMFS-TM-AFSC-485

May 2024

Abstract

NOAA's Alaska Fisheries Science Center's (AFSC) Ecosystems and Fisheries-Oceanography Coordinated Investigations (EcoFOCI) program has collected spring ichthyoplankton abundance data in the Gulf of Alaska since 1981. Collections were made nearly annually until 2011 when sampling was reduced to only odd years. This dataset is used to better understand population recruitment of major fish species in the GOA and provides early warning of potential year-class strength to inform fisheries management. However, gaps in the time series during even years have made it more difficult to interpret the interannual variability of ichthyoplankton abundance in such a dynamic ecosystem. Recent collaboration with the Northern Gulf of Alaska Long Term Ecological Research (NGA LTER) program has allowed for additional spring sampling of ichthyoplankton in the GOA annually since 2018. Larval fish data collected by the NGA LTER were combined with EcoFOCI data and used to estimate abundance in years when EcoFOCI had no field presence in the GOA. Five taxa were determined to be suitable for this approach based on their percent occurrence in both surveys. A generalized additive model (GAM) was fit to ichthyoplankton data from 1981 to 2022 collected by both EcoFOCI and NGA LTER and used to predict larval abundances in 2018, 2020, and 2022. For each species, models with two different error distributions were compared and shown to produce similar predictions of larval abundance. This report provides a model framework for predicting interannual larval fish abundance while controlling for differences in sampling methodologies, timing, and location, and identifies a subset of taxa for which this framework is currently appropriate. As additional years of concurrent sampling are added in future, this approach has the potential to improve our understanding of interannual variation in ichthyoplankton dynamics and provide more comprehensive indicators for ecosystem-based fisheries management.

Contents

Introduction

The Gulf of Alaska (GOA) is a highly productive ecosystem which supports many commercially and ecologically important fish. NOAA's Alaska Fisheries Science Center is responsible for monitoring and providing management advice for these important fisheries. Ecosystem indicators, including climate indices, biomass of prey and predator communities, and socioeconomic conditions are an important component to assessing the status of a stock and informing set catch limits through ecosystem-based fisheries management (Zador et al. 2017). Larval fish abundances serve as a valuable indicator of the potential recruitment strength for a given year class (Bailey et al. 2012). The larval stage is particularly sensitive to environmental changes; thus, changes in the population abundance, distribution, and phenology can provide insight into current spawning, habitat, and forage conditions (Boeing and Duffy-Anderson 2008, Asch 2015, Auth et al. 2018, Rogers and Dougherty 2019, Nielsen et al. 2021, Rogers et al. 2021). Tracking abundances of larval fish allows for early insight into the potential future status of a stock and has been used to inform catch limits (Litzow et al. 2022).

The Ecosystem and Fisheries-Oceanography Coordinated Investigations (EcoFOCI) team at NOAA conducts regular monitoring of spring ichthyoplankton communities in the Gulf of Alaska during May - June. Larval fish abundances in the Gulf of Alaska peak during the spring bloom when phytoplankton and zooplankton biomass are high and the continental shelf provides important habitat for many species (Doyle et al. 2019). The EcoFOCI ichthyoplankton survey was historically designed to target the spring peak in larval fish abundance with particular focus on the commercially important walleye pollock (*Gadus chalcogrammus*). Despite the initial focus on walleye pollock, all larval fish are identified and enumerated within a sample, providing a rich dataset on interannual ichthyoplankton abundances for many important fish species in the Gulf of Alaska region (Matarese et al. 2003). Time series of larval abundance have been estimated for a subset of 12 commercially and ecologically important taxa and contributed as indicators to annual Ecosystem Status Reports (ESRs; Rogers and Axler 2023).

Samples are collected primarily in the western Gulf of Alaska with the most consistent sampling occurring in Shelikof Strait (Fig. 1), one of the primary spawning grounds of walleye pollock. EcoFOCI has collected and processed ichthyoplankton samples since 1981. Collections were made annually from 1990 until 2011 when the program was reduced to sampling every other year. Gaps in the time series have made it more difficult to interpret the temporal variability of ichthyoplankton abundance particularly in recent years which have experienced increasing environmental variability and marine heatwave events (Nielsen et al. 2021, Suryan et al. 2021, Ren et al. 2023). In years without sampling, a lack of observations prevents detecting early warning signals of failed recruitment or large-scale ecosystem shifts.

Recent collaboration with the Northern Gulf of Alaska Long Term Ecological Research (NGA LTER) program has allowed for additional collections of ichthyoplankton. The NGA LTER was established in 2018 as a part of a network of 30 NSF-funded long-term ecological research programs and one of four pelagic-focused programs. It builds off over 30 years of prior timeseries collections in the Gulf of Alaska. The NGA LTER began collecting additional ichthyoplankton samples during their annual spring cruise in collaboration with EcoFOCI. Despite the difference in timing and location of collections between the NGA LTER and EcoFOCI, NGA LTER collections could potentially be used to predict larval fish abundances in years without EcoFOCI sampling and help inform interannual patterns in abundance during sampling gaps. In this study, we investigated model-based approaches for combining the EcoFOCI and NGA LTER ichthyoplankton datasets to estimate time-series of larval fish abundance for a subset of commercially and ecologically important species in the Gulf of Alaska.

Methods

Ichthyoplankton Collections

EcoFOCI has collected ichthyoplankton samples from the Western Gulf of Alaska between May and June since 1981 (Fig. 2). Sample locations span the western Gulf of Alaska shelf with the area between Shelikof Strait and the Shumagin Islands being most consistently sampled (Fig. 1). Sample coverage varied by year (Table 2; Fig. 3). Samples were taken annually between 1990 and 2011, after which sampling was reduced to every other year on odd years (Table 2). A 60 cm diameter bongo net equipped with either 333 or 505 µm mesh nets was towed obliquely from 100 m depth, or 10 m off bottom, to the surface (Table 1). Larval catch has been shown to be comparable between the two mesh sizes (Boeing and Duffy-Anderson 2008). Collections were made during both night and day as ichthyoplankton are assumed to remain in the upper 100 m (Brodeur and Rugen 1994). Samples were preserved in 5% formalin, identified to the lowest taxonomic level at the Plankton Sorting and Identification Center in Szczecin, Poland, and verified by taxonomic experts at the Alaska Fisheries Science Center. Larval catch is reported as the number per 10 m^2 sea surface area.

The NGA LTER began collecting ichthyoplankton annually for EcoFOCI in 2018 (Table 2). Samples are collected each spring between April and May along three core cross-shelf transects in the NGA LTER study region (Figs. 1, 2). Additional stations were sampled as time permitted (Fig. 3). Either a 60 cm bongo net or a 0.25 m^2 Hydro-Bios MultiNet with a drogue net were towed obliquely from 200 m depth, or 5 m above bottom, to the surface at night. Both net systems were equipped with 505 µm mesh nets (Table 1). Samples were preserved in 10% formalin and identification was carried out following EcoFOCI protocols as described above.

A total of 4,823 spring ichthyoplankton collections were made over 35 years between 1981 and 2022. The number, timing, and location of collections varied by year, with EcoFOCI and NGA LTER making an average of 146 ± 64 and 30 ± 9 collections per year, respectively (Table 2, Fig. 3). On average, EcoFOCI collections were made 26 days later than NGA LTER collections (Fig. 2). The majority of EcoFOCI sample locations were on the continental shelf with a bottom depth shallower than 500 m, whereas NGA LTER samples transected the shelf with some samples over the continental slope. 2019 and 2021 are the only years sampled by both EcoFOCI and NGA LTER in the same year.

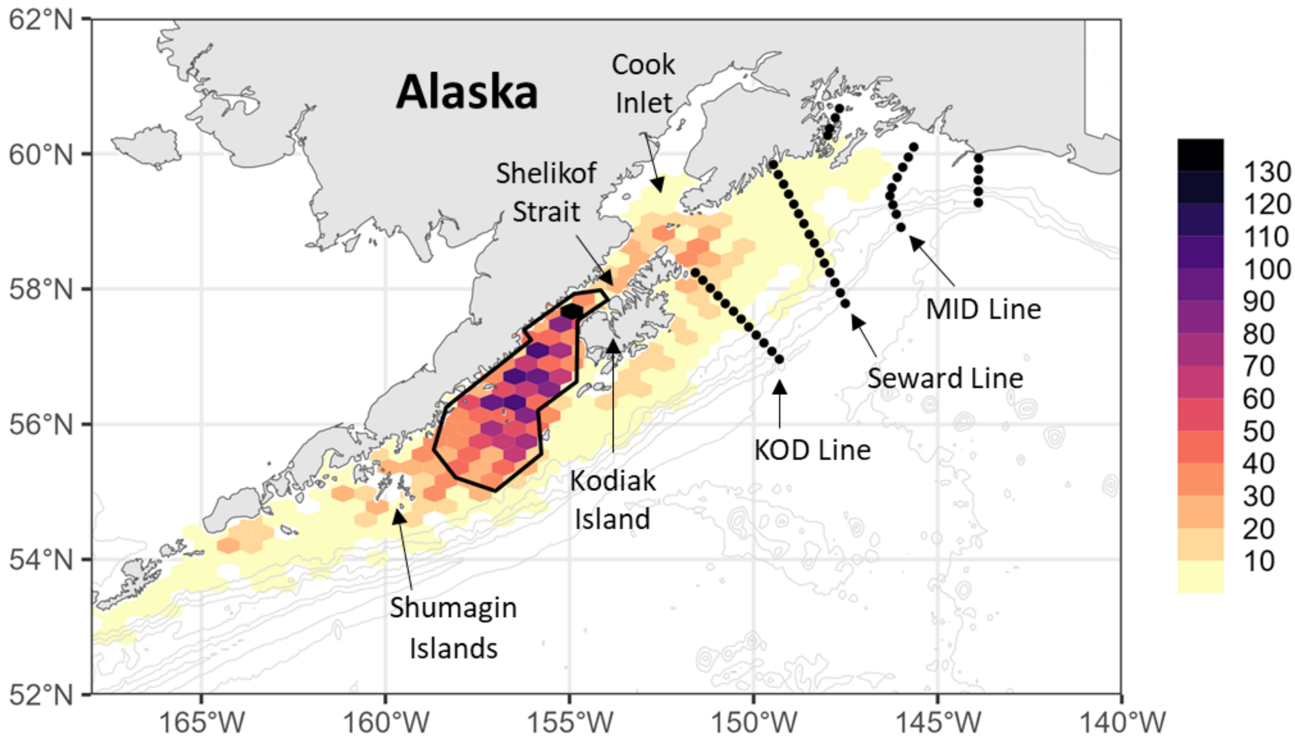

Figure 1.-- Map of the Gulf of Alaska sample region. Color indicates the number of samples collected by EcoFOCI within the hexagonal grid from 1981 to 2022. NGA LTER sample locations are marked by points and main transect lines are labeled. The 1,000 m isobath contours are drawn in gray. Black polygon outline indicates the region used to generate predicted abundances.

Table 1. -- Summary of collection conditions for EcoFOCI and NGA LTER.

Group	Equipment	Mesh Size	Max Tow Depth	Bottom Depth	Time of Tow	Longitude Range	Date Range
EcoFOCI	60 cm Bongo Net	333 or 505 µm	100 m	18 – 3580 m (avg = 179 m)	Day and Night	146.50 to 168.00°W	6 May to 6 June
NGA LTER	60 cm Bongo Net or 0.25 m^2 Mulitnet	505 µm	200 m	35 – 4548 m (avg = 830 m)	Night	143.89 to 151.59°W	20 April to 10 May

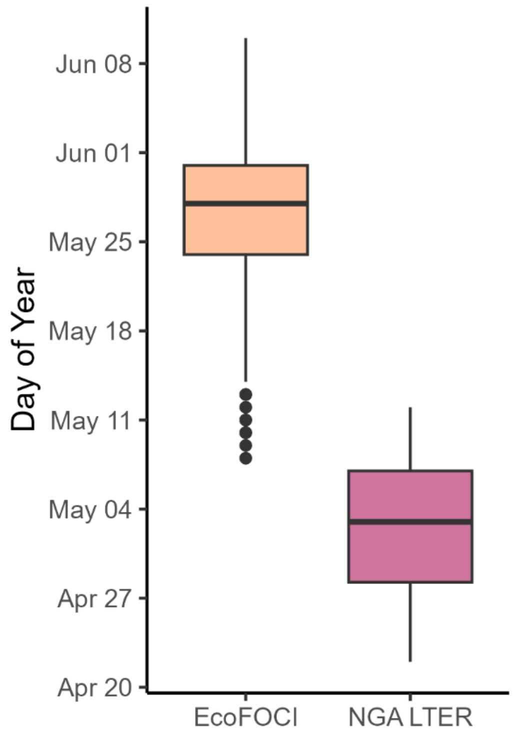

Figure 2. -- Boxplot of dates sampled for EcoFOCI and NGA LTER.

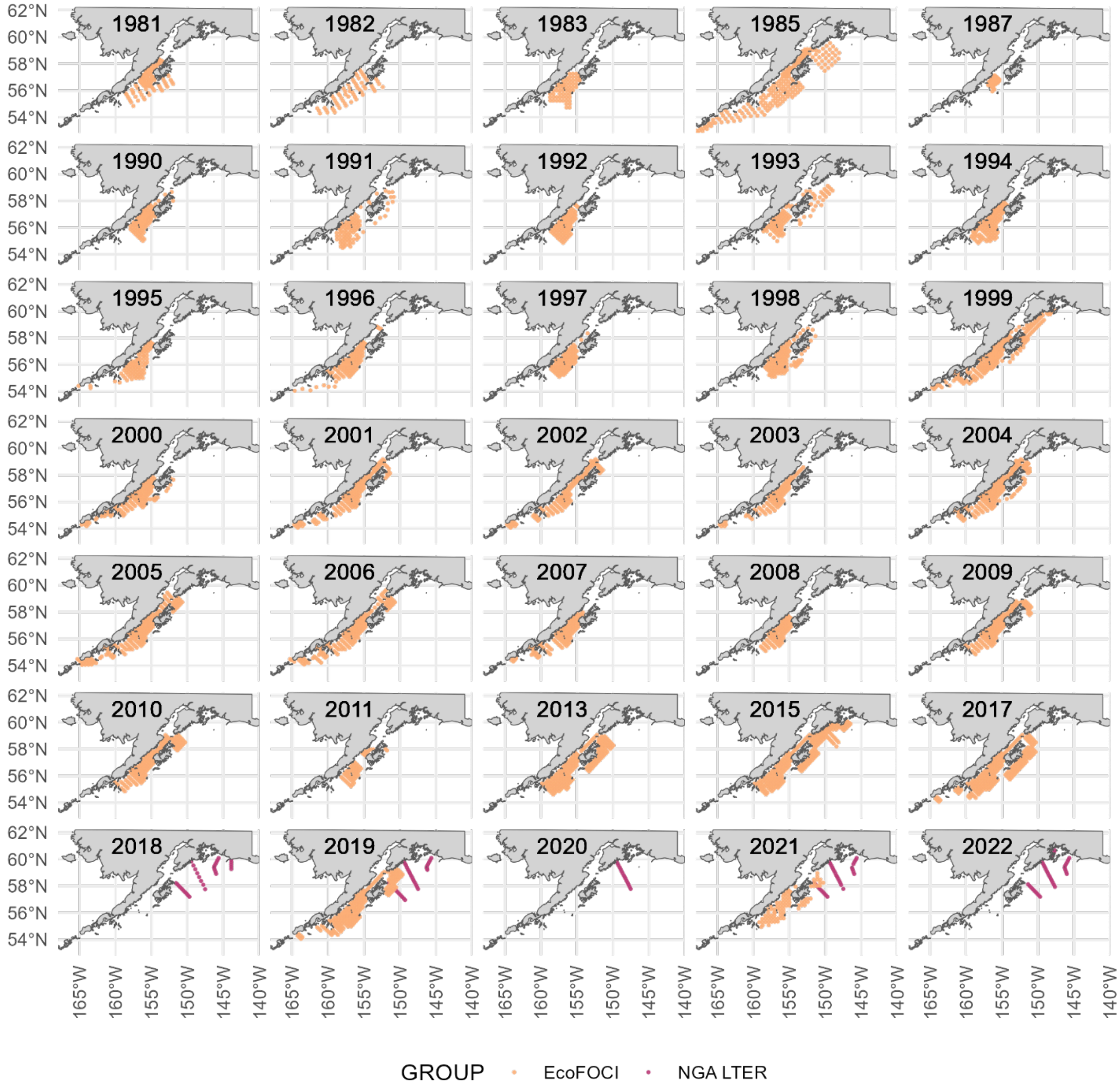

Figure 3. -- Map of ichthyoplankton sample locations per year. Colors indicate EcoFOCI (orange) or NGA LTER (purple) sampling.

Table 2. -- Number of samples collected per year for EcoFOCI and NGA LTER since 1981.

Year	EcoFOCI	NGA LTER	Year	EcoFOCI	NGA LTER
1981	130	0	2003	114	0
1982	61	0	2004	193	0
1983	69	0	2005	187	0
1984	0	0	2006	175	0
1985	185	0	2007	130	0
1986	0	0	2008	94	0
1987	50	0	2009	152	0
1988	0	0	2010	163	0
1989	0	0	2011	66	0
1990	133	0	2012	0	0
1991	96	0	2013	226	0
1992	137	0	2014	0	0
1993	113	0	2015	281	0
1994	139	0	2016	0	0
1995	98	0	2017	266	0
1996	130	0	2018	0	32
1997	100	0	2019	232	35
1998	130	0	2020	0	15
1999	314	0	2021	84	32
2000	141	0	2022	0	36
2001	147	0	Total	4,673	150
2002	137	0			

Larval Catch Models

Generalized additive models (GAM) were used to estimate interannual patterns in larval fish abundance and control for differences in spatial sampling and methodologies between EcoFOCI and NGA LTER collections. Species included in EcoFOCI's ESR contributions were considered for this model approach due to their high abundance and importance in the Gulf of Alaska. These included Pacific sand lance (*Ammodytes personatus*), arrowtooth flounder (*Atheresthes stomias*), ronquils (*Bathymaster* spp.), walleye pollock (*Gadus chalcogrammus*), Pacific cod (*Gadus macrocephalus*), flathead sole (*Hippoglossoides elassodon*), Pacific halibut (*Hippoglossus stenolepis*), southern rock sole (*Lepidopsetta bilineata*), northern rock sole (*Lepidopsetta polyxystra*), starry flounder (*Platichthys stellatus*), rockfish (*Sebastes* spp.), and northern lampfish (*Stenobrachius* spp). Due to low occurrence in NGA LTER samples, some species were deemed unsuitable for model generated predictions. Species not present in at least 30% of NGA LTER samples were eliminated. Years with no catch for a given species at any location were excluded from the model for a given species. To account for the highly zero-inflated data structure, two types of models were tested, which differed in their approach to handling zeros.

First, a Tweedie distribution with a log-link was selected for its ability to model continuous data with a point mass at 0 (Dunn and Smyth 2018). The Tweedie power parameter was estimated within the model. For each species, two alternative Tweedie models were fit. The first modeled catch per 10 m^2 (CPUE) as a function of Year (as a factor) and a two-dimensional spatial smoothers (Easting, Northing). Longitude and latitude coordinates were converted to UTM zone 5 coordinates prior to input into the model. The second model contained an additional term for research Group (EcoFOCI or NGA LTER) to further account for differences in collection methods and timing (Equation 1). Model comparison using AIC showed that, for most species, the inclusion of a research group term resulted in a better fit. All models discussed hereafter include a research Group fixed effect (Equation 1). Models with terms for day of year, sample net, time of day, and maximum tow depth were not considered due to collinearity and were assumed to be accounted for in the group term. No interaction terms between Year, Group, and spatial smooth were included due to limited data overlap:

$$log(CPUE) = s(Easting, Northing) + Year + Group. \qquad \text{Eq. 1}$$

The second model distribution used was a two-part hurdle model (also known as a delta model), selected for its flexibility to model extra zeros separate from positive catch data. First, larval presence and absence (p_i) was modeled with a binomial distribution with a logit link (Equation 2). Then non-zero CPUE (μ_i) was modeled with a Gamma log-link distribution (Equation 3). Models were parameterized as described for the Tweedie model (Equation 1). All models were constructed using the mgcv package (v1.9.0; Wood 2017) and visualized with the gratia (v0.8.1; Simpson 2023) package in R (v4.3.1; R core team 2023):

$$logit(p_i) = s(Easting, Northing) + Year + Group \qquad \text{Eq. 2}$$
$$log(\mu_i) = s(Easting, Northing) + Year + Group. \qquad \text{Eq. 3}$$

Log likelihood and AIC values for the Tweedie and Hurdle models were used to compare model fit (Zhen et al. 2018). For the Hurdle model, log likelihood was calculated as the sum of the log likelihoods from the binomial and nonzero parts (McDowell 2003). AIC was then calculated as AIC = -2logL + 2p where logL is the log likelihood and p is the sum of the model degrees of freedom.

Predicted Abundances

A prediction grid with 12 km spacing was generated for the region typically reported in ESRs between Shelikof Strait and the Shumagin islands (Fig. 1, polygon). For each model type, CPUE was predicted at each point within the prediction grid for every sampled year, including years when only NGA LTER stations were sampled. The Group factor was set to EcoFOCI. For the hurdle model, binomial and nonzero model predictions were multiplied to generate an overall predicted CPUE for each point on the grid. To create a timeseries, predicted abundances for the grid were then averaged per year. Two-sided nonparametric bootstrapped 95% confidence intervals were generated for annual predicted CPUE. Data was resampled randomly with replacement then used to generate predicted abundances for 1,000 replicates. Confidence intervals were calculated using the bootstrap percentile method. For the hurdle model, random resampling resulted in all positive data for a given year being dropped for some replicates, leading to an error. In this case, the larval catch value for one random sample in that year was replaced with a pseudocount of 0.001-- representing a near-zero value while allowing the function to proceed without error. Bootstrapped confidence intervals were generated using the boot package in R (v1.3.28.1; Canty and Ripley 2022, Davison and Hinkley 1997).

Model Prediction Testing

Two years (2019 and 2021) were sampled by both EcoFOCI and the NGA LTER. To test the model's ability to predict abundances for EcoFOCI samples based on NGA LTER collections, all models were refit leaving out either 2019 or 2021 EcoFOCI data and compared to the full model. New predicted values were calculated for CPUE as described above. Root mean squared error (RMSE) was calculated between the predictions generated by leaving out 2019 or 2021 and the full model predictions for that year. RMSE was calculated for both Tweedie and Gamma Hurdle models of each species to assess which model type performed better when removing a year of data per species.

Results

Of the 12 species included in ESRs, 5 were present in at least 30% of NGA LTER samples and thus were modeled (arrowtooth flounder, northern lampfish, Pacific sand lance, rockfish, and walleye pollock). The other seven species were not modeled due to a lack of presence in NGA LTER collections and thus low confidence in the model's predictive capabilities for that species (Table 3). Of the five species, walleye pollock was the most ubiquitous with presence in 83% of EcoFOCI samples and 65% of NGA LTER samples. Pacific sand lance also had a higher percent occurrence in EcoFOCI collections. Arrowtooth flounder, rockfish, and northern lampfish had higher percent occurrence in NGA LTER samples (Table 3).

Table 3. --Percentage of samples a species is present for EcoFOCI and NGA LTER collections. Model predicted abundances were calculated for species with percent occurrence greater than 30% in NGA LTER samples (bolded).

Common Name	Scientific Name	% Occurrence EcoFOCI	% Occurrence NGA LTER
Pacific sand lance	***Ammodytes personatus***	**71.5**	**36.7**
Arrowtooth flounder	***Atheresthes stomias***	**19.9**	**46.7**
Ronquils	*Bathymaster* spp.	59.7	4.7
Walleye pollock	***Gadus chalcogrammus***	**82.8**	**64.7**
Pacific cod	*Gadus macrocephalus*	43.8	9.3
Flathead sole	*Hippoglossoides elassodon*	75.0	20.0
Pacific halibut	*Hippoglossus stenolepis*	9.3	8.0
Southern rock sole	*Lepidopsetta bilineata*	27.4	12.0
Northern rock sole	*Lepidopsetta polyxystra*	32.5	8.7
Starry flounder	*Platichthys stellatus*	17.5	3.3
Rockfish	***Sebastes* spp.**	**43.1**	**72.0**
Northern lampfish	***Stenobrachius* spp.**	**38.8**	**56.7**

Pacific sand lance, arrowtooth flounder, and northern lampfish were best modeled by the Tweedie distribution, as indicated by a lower AIC value (Table 4). Conversely, walleye pollock and rockfish were best modeled by the Gamma Hurdle model (Table 4). However, model estimates were very similar between Tweedie and Gamma Hurdle models. Both Tweedie and Gamma Hurdle model residuals were inspected and deemed acceptable.

Partial effects for the two-dimensional smoother displayed similar spatial patterns in abundance (Figs. 4-6). Arrowtooth flounder, northern lampfish, and rockfish had higher presence and abundances off shelf. Additionally, northern lampfish and rockfish showed higher abundances on the shelf near the Seward line transect. Walleye pollock and Pacific sand lance had higher larval catch on shelf with walleye pollock more abundant in the Shelikof Strait region and Pacific sand lance more abundant nearshore. It is important to note that the spatial smoother does not purely represent spatial patterns in larval catch between NGA LTER and EcoFOCI due to the spatial collinearity with sample timing and methodological differences. Thus, patterns represented by the spatial smoother reflect the combined impact of spatial and temporal patterns on larval abundance, making it a useful approach for controlling for differences in larval catch as a result of both spatial distributions and methodological differences.

Table 4. --Log likelihood and associated degrees of freedom (df), Akaike information criterion (AIC), and root mean squared error (RMSE) for Tweedie and Gamma Hurdle models for each species. The model with the lowest AIC and RMSE error per species is highlighted in gray indicating better model fit.

	Log Likelihood		AIC		RMSE	
Species	Tweedie	Hurdle	Tweedie	Hurdle	Tweedie	Hurdle
Arrowtooth flounder	-5845.79 (df=64.8)	-5807.65 (df=125.1)	11821.18	11865.58	6.01	7.20
Northern lampfish	-9696.77 (df=65.4)	-9771.13 (df=123.7)	19524.35	19789.68	10.92	11.67
Pacific sand lance	-19022.88 (df=65.6)	-19293.49 (df=124.9)	38176.92	38836.86	105.12	98.98
Rockfish	-13698.74 (df=65.8)	-13487.92 (df=126.4)	27529.07	27228.58	52.74	28.49
Walleye pollock	-26480.93 (df=65.7)	-26160.88 (df=126.7)	53093.25	52575.22	143.86	73.85

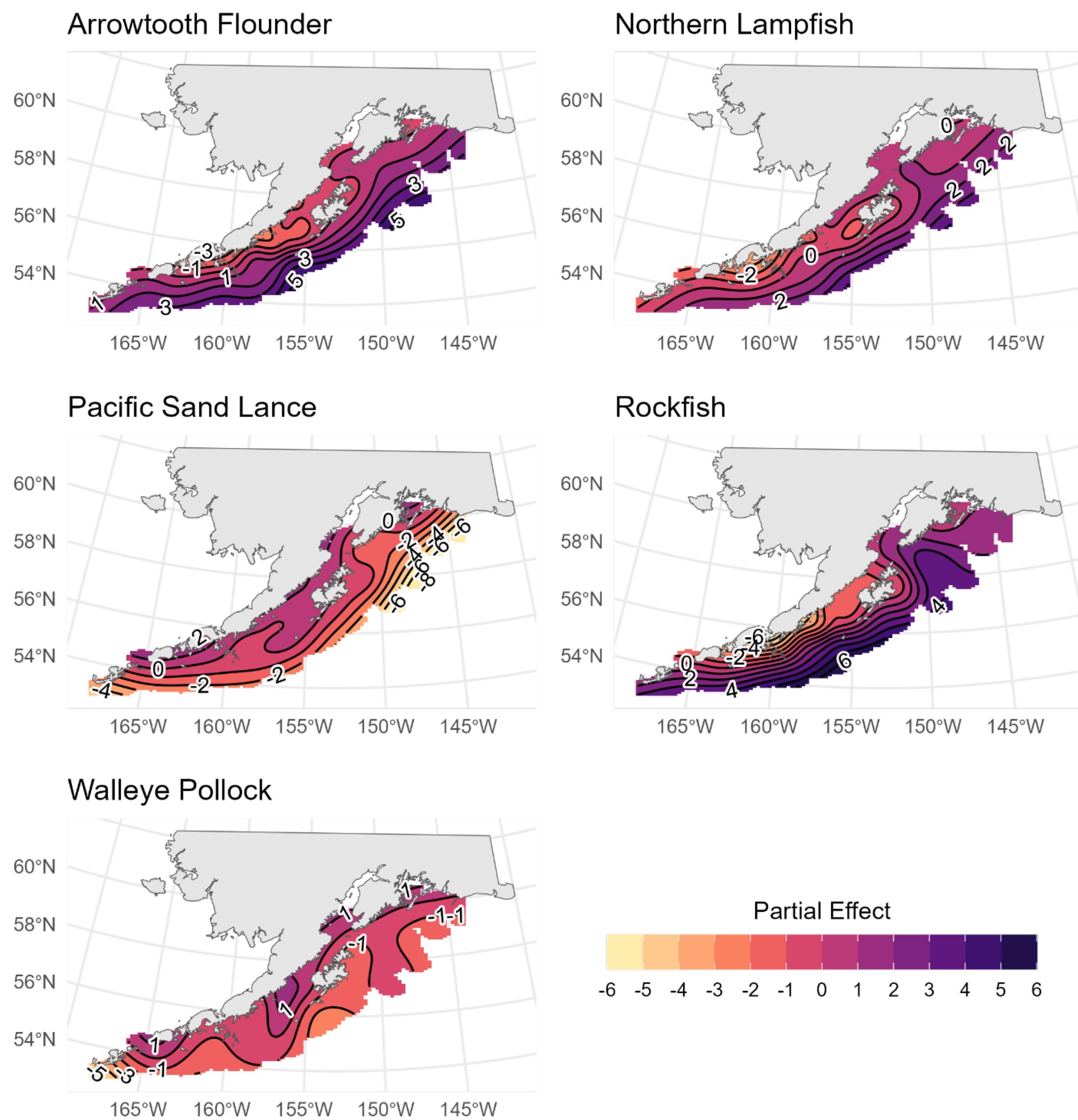

Figure 4. -- Tweedie model two-dimensional spatial smoother partial effects. Positive values indicate CPUE greater than average and negative values indicate CPUE lower than average.

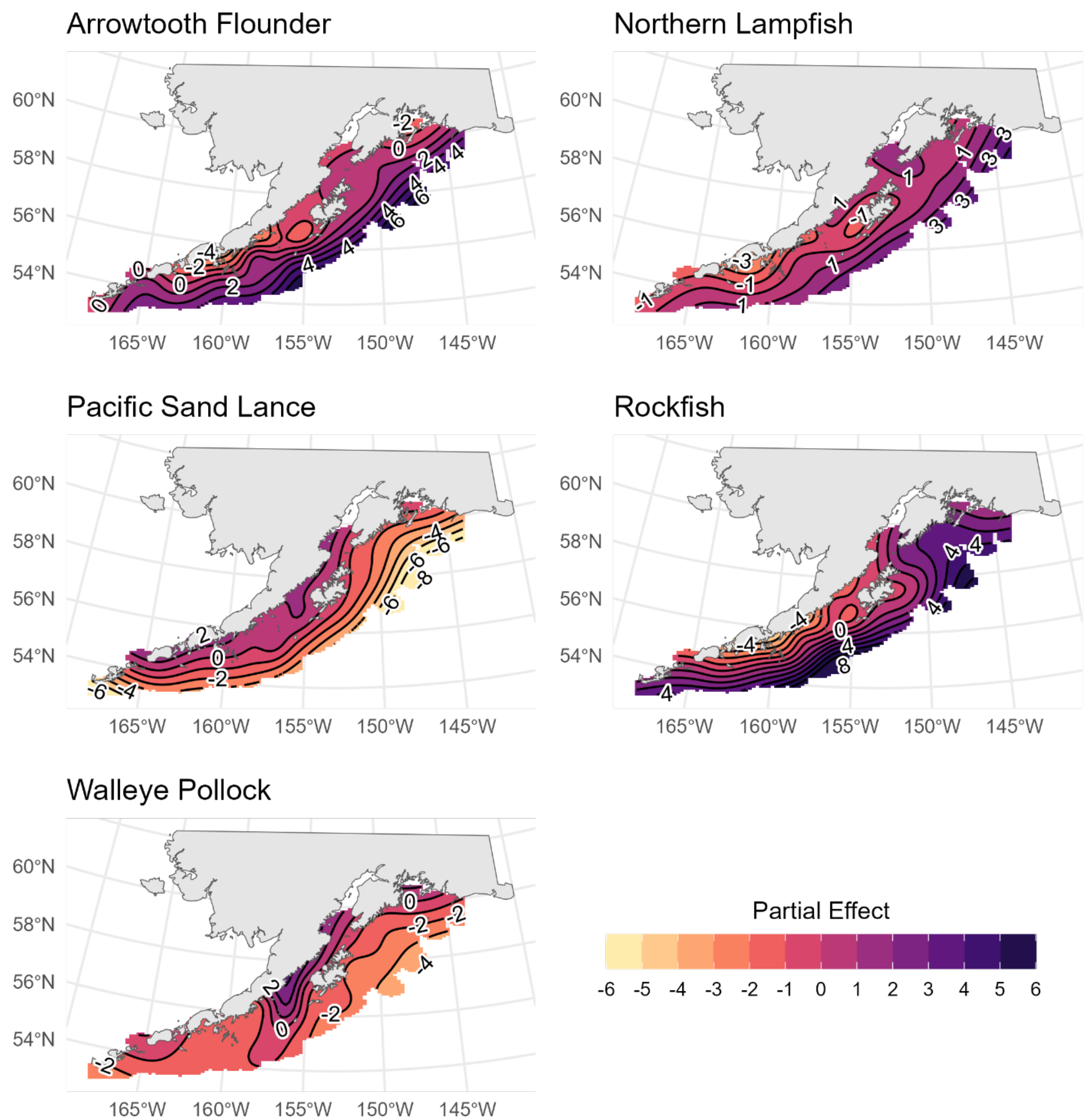

Figure 5. -- Gamma Hurdle binomial model two-dimensional spatial smoother partial effects. Positive values indicate higher than average likelihood of positive catch and negative values indicate lower than average likelihood of positive catch.

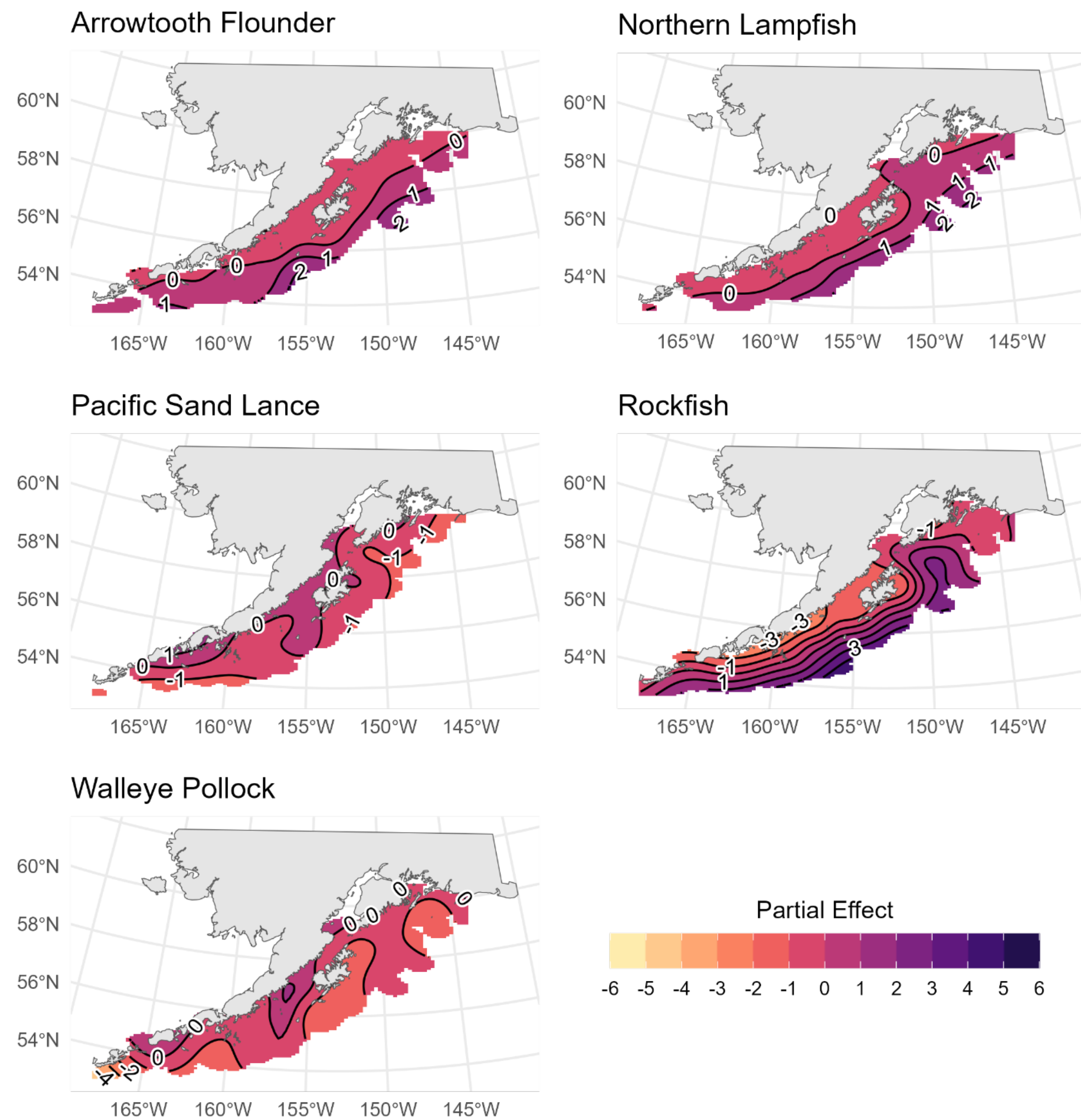

Figure 6. -- Non-zero Gamma Hurdle model two-dimensional spatial smoother partial effects. Positive values indicate higher than average CPUE when larvae are present and negative values indicate lower than average CPUE when larvae are present.

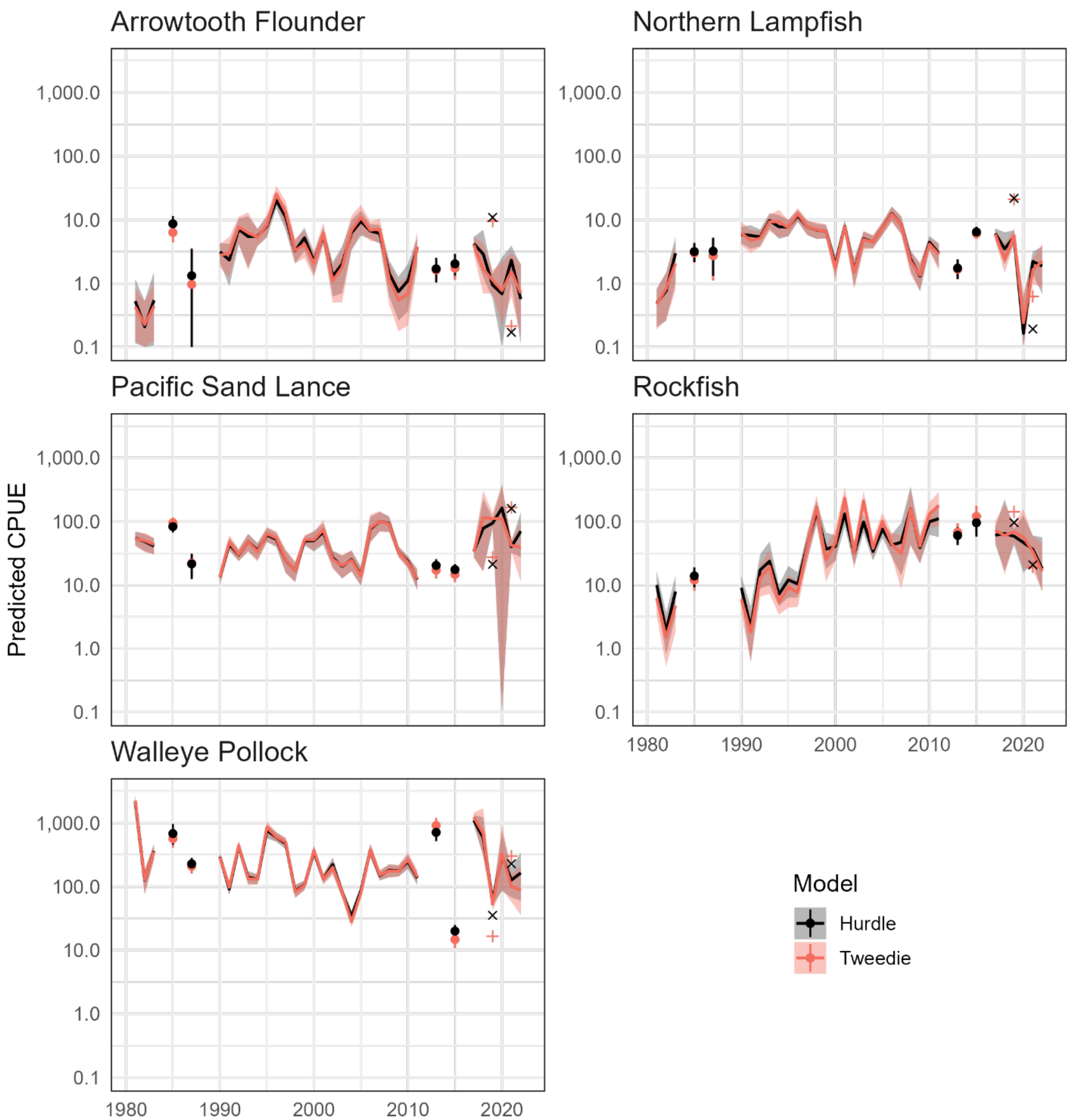

Figure 7. -- Model predicted catch per unit effort (CPUE; Larvae per 10 m^2) per year per species. Data are plotted on the log10 scale. Gamma Hurdle model predictions are plotted in black and Tweedie model predictions are plotted in pink. Nonparametric bootstrapped 95% confidence intervals are represented by shading and error bars. Black × and pink + symbols indicate predictions for Hurdle and Tweedie models, respectively, based on the removal of 2019 or 2021 EcoFOCI data.

There is substantial overlap of model predicted CPUE and associated confidence intervals between Tweedie and Gamma Hurdle models (Fig. 7). Each species showed high interannual variability, often exceeding one order of magnitude difference. Conversely, variability within a year was typically smaller. Years predicted solely based on NGA LTER collections have larger confidence intervals than other years. Additionally, for some species, some years have particularly large confidence intervals due to lower sample sizes or restricted spatial coverage. Over the sample period, walleye pollock has the highest predicted CPUE, Pacific sand lance and rockfish have intermediate CPUE, and arrowtooth flounder and northern lampfish have the lowest CPUE.

Model predictions for 2019 and 2021 generated based on models built without 2019 or 2021 EcoFOCI data, respectively, did not fall within the 95% confidence intervals of the full model predictions (Fig. 7). The exception is rockfish with 2021 EcoFOCI data removed. For walleye pollock and Pacific sand lance, predictions were typically lower for models with 2019 EcoFOCI data removed and higher for models built without 2021 data compared to the full model. The opposite pattern is true for arrowtooth flounder, northern lampfish, and rockfish (Fig. 7). In most cases, the Tweedie and Gamma Hurdle model performed similarly with the removal of EcoFOCI data. Arrowtooth flounder and northern lampfish had lower RMSE with the Tweedie model, whereas Pacific sand lance, walleye pollock, and rockfish had lower RMSE with the Gamma Hurdle model (Table 4). Preference for Tweedie or Gamma Hurdle model distributions based on RMSE was not uniform across modeled species for the current model framework and available data.

Discussion

The aim of this analysis was to utilize annual ichthyoplankton collections made by the NGA LTER program to predict larval fish abundances during years which EcoFOCI did not make collections. Here, we outline a simple model framework for predicting annual abundances of five key fish species in the Gulf of Alaska based on NGA LTER collections while controlling for spatial, temporal, and methodological differences in sample collections. Although this method is currently limited to two years of overlapping data, it provides a framework for continuing to predict abundances as future sample data becomes available with potential for extension to other important species.

Due to the NGA LTER's earlier collection timing, this method works best for species with earlier life history phenologies. Spawning and hatching occurs prior to March for the five species included, allowing for accumulation of larvae leading up to the early NGA LTER collections. Conversely, for the excluded species, spawning does not occur until April and peaks later (Doyle et al. 2019). Thus, NGA LTER collections are likely too early to collect consistent numbers of these species. The exception is the winter spawned *Hippoglossus stenolepis*, which is generally patchy and present in low numbers.

Data collected east of the Cook Inlet was generated by the NGA LTER (with the exception of 1985, 1999, and 2015 when EcoFOCI sampling extended farther east) and subject to differences in methodologies and earlier sample timing (Fig. 3). Thus, patterns in the two-dimensional smoother across these regions reflect not only spatial patterns in larval abundance, but also potential changes due to differential sample timing and collection methods. It is important to not interpret the two-dimensional partial effects purely as a spatial term even though it is defined mathematically as such. For example, *Sebastes* spp., and *Stenobrachius* spp. show a region of higher abundances that crosses the shelf near the Seward Line (Fig. 7). This may represent a true spatial pattern reflective of the transport of deep-spawned larvae onto the shelf via the Amatuli Trough (Doyle et al. 2019). However, this cannot be separated from the confounding impact of the NGA LTER's earlier sample timing and potential increase in catchability due to the deeper tow depth and use of the Multinet on the Seward Line. Additionally, the two-dimensional smooth parameter was not allowed to fluctuate per year and therefore assumes that these patterns remain constant over time. The addition of the research group parameter to the model framework adds additional flexibility to account for differences between NGA LTER and EcoFOCI collections.

For each species, predicted CPUE had high interannual variability over the sample period, often exceeding one order of magnitude (Fig. 7). Conversely, the prediction error due to variability across sampling locations and time frames was typically smaller. In some cases, restricted sampling range and sample quantity resulted in some years with low percent occurrence for some species and large confidence intervals. For example, fewer samples were taken in 1987 and restricted to the shelf resulting in low catch of arrowtooth flounder which is more abundant offshore. Likewise, sample collection was greatly reduced in 2020 with only 15 samples collected due to the COVID-19 pandemic (Fig. 3, Table 2). In these years, bootstrapped confidence intervals are large due to the higher probability of random resampling selecting data with a CPUE of zero (Fig. 7). However, overall, the greater interannual variability in comparison to spatial and within-year variability makes this modeling approach useful to approximate annual larval catch from NGA LTER collections despite differences in sampling location and timing.

Model predictions for both the Tweedie and Gamma Hurdle models estimated similar absolute CPUE, relative interannual patterns, and confidence intervals (Fig. 7). Models using Tweedie and Gamma Hurdle methods have been recommended previously on Bering Sea and Gulf of Alaska bottom trawl fish abundance time series and shown to produce similar scale estimates to each other. Moreover, Tweedie and Gamma Hurdle GAMs were shown to produce similar absolute values to the design-based indices used compared to inverse Gaussian or lognormal distributions (Thorson et al. 2021). While the absolute abundance of larval fish in the EcoFOCI sample region is unknown for 2018, 2020, and 2022, both model types produce similar CPUE predictions relative to the surrounding sampled years. Predicting the absolute CPUE could be useful; however, correctly modeling relative interannual changes in CPUE was the primary purpose in our analysis.

Models were subjected to rigorous testing by leaving out EcoFOCI data for one year of overlapping sampling and regenerating model predicted abundances for that year on the reduced dataset. Resulting CPUE predictions did not fall within the 95% confidence intervals of the full model, illustrating the model's sensitivity to the removal of data. Based on RMSE calculations, the Gamma Hurdle model performed better on the reduced data set for Pacific sand lance, rockfish, and walleye pollock. Conversely, the Tweedie model performed better on arrowtooth flounder and northern lampfish. Because the current dataset only contains two years that were sampled concurrently by EcoFOCI and the NGA LTER (2019 and 2021), the removal of one year resulted in models being constructed with only one year of overlap. It is expected that the addition of more years of overlapping sampling in the future will help stabilize model predictions and enhance its reliability. It is unclear how the addition of more data will impact the selection between Tweedie or Gamma Hurdle models.

For walleye pollock in particular, larval abundance estimates have been used as an early indicator of potential year-class strength (Bailey et al. 2012, Litzow et al. 2022). However, recent time-series gaps have limited the utility of this indicator for informing the stock assessment for walleye pollock (Shotwell et al. 2019). Despite the limited data in 2018, 2020, and 2022, relative patterns in model-predicted CPUE for those years are consistent with eventual recruitment strength as estimated in the stock assessment model (Monnahan et al. 2023). In the assessment model, which estimates recruitment strength based on survey and fishery catches of walleye pollock aged 1 and older, recent strong year classes have been 2017, 2018, and 2020, while 2015, 2019, 2021 and 2022 were estimated to be weak. This corresponds to years with relatively high and low model-estimated larval abundances (Fig. 7). This suggests that even with limited sampling that occurs outside the typical time range and geographic area of the EcoFOCI survey, the NGA LTER survey provides useful estimates of relative abundance when modeled in this way. In the future, a combined time series can be contributed to the Ecosystem and Socioeconomic Profile for this stock (Shotwell et al. 2019).

Current and future climate changes are expected to impact ecosystems and fisheries production and there is an increasing need for ecosystem monitoring. Moreover, climate change is expected to increase environmental variability and frequency of extreme events like marine heatwaves, necessitating frequent sampling efforts to capture variability on shorter time scales (Ren et al. 2023). Nonetheless, limitations to funding and ship time often curtail the scope and frequency of sampling efforts, thereby increasing the importance of partnerships to fill in information and data gaps. Thus, development of quantitative methodology for combining data sources across multiple sampling platforms and drawing inferences becomes necessary. This report provides a simple model framework for predicting interannual larval fish abundance while controlling for differences in sampling methodologies, timing, and location, and identifies a subset of taxa for which this framework is currently appropriate. As additional years of concurrent sampling are added in future, this approach has the potential to improve our understanding of interannual variation in ichthyoplankton dynamics and provide more comprehensive indicators for ecosystem-based fisheries management.

Citations

Asch, R. G. 2015. Climate change and decadal shifts in the phenology of larval fishes in the California Current ecosystem. Proceedings of the National Academy of Sciences 112(30):E4065–E4074.

Auth, T. D., E. A. Daly, R. D. Brodeur, and J. L. Fisher. 2018. Phenological and distributional shifts in ichthyoplankton associated with recent warming in the northeast Pacific Ocean. Global Change Biology 24(1):259–272.

Bailey, K., T. Zhang, K. Chan, S. Porter, and A. Dougherty. 2012. Near real-time forecasting of recruitment from larval surveys: Application to Alaska pollock. Marine Ecology Progress Series 452:205–217.

Boeing, W. J., and J. T. Duffy-Anderson. 2008. Ichthyoplankton dynamics and biodiversity in the Gulf of Alaska: Responses to environmental change. Ecological Indicators 8(3):292–302.

Brodeur, R., and W. C. Rugen. 1994. Diel vertical distribution of ichthyoplankton in the northern Gulf of Alaska. Fishery Bulletin, U.S. 92:223–235.

Canty, A., and B. Ripley. 2022. boot: Bootstrap R (S-Plus) Functions. R package version 1.3-28.1.

Davison, A. C., and D. V . Hinkley. 1997. Bootstrap Methods and Their Applications. Cambridge University Press, Cambridge. ISBN 0-521-57391-2

Doyle, M. J., S. L. Strom, K. O. Coyle, A. J. Hermann, C. Ladd, A. C. Matarese, S. K. Shotwell, and R. R. Hopcroft. 2019. Early life history phenology among Gulf of Alaska fish species: Strategies, synchronies, and sensitivities. Deep Sea Research Part II: Topical Studies in Oceanography 165:41–73.

Dunn, P. K., and G. K. Smyth. 2018. Generalized Linear Models With Examples in R. Springer New York, New York, NY.

Litzow, M. A., A. A. Abookire, J. T. Duffy-Anderson, B. J. Laurel, M. J. Malick, and L. A. Rogers. 2022. Predicting year class strength for climate-stressed gadid stocks in the Gulf of Alaska. Fisheries Research 249:106250.

Matarese, A. C., D. M. Blood, S. J. Picquelle, and J. L. Benson. 2003. Atlas of abundance and distribution patterns of ichthyoplankton from the northeast Pacific Ocean and Bering Sea ecosystems based on research conducted by the Alaska Fisheries Science Center (1972-1996). U.S. Department of Commerce, NOAA Professional Papers NMFS-SPO-1, 281 p.

McDowell, A. 2003. From the Help Desk: Hurdle Models. The Stata Journal 3(2):178–184.

Monnahan, C., G. D. Adams, B. E. Ferriss, S. K. Shotwell, D. R. McKelvey, and D. McGowan. 2023. Assessment of the walleye pollock Stock in the Gulf of Alaska, p. 1-124. *In* Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska for 2023. North Pacific Fishery Management Council, 1007 West Third, Suite 400, Anchorage, AK 99501.

Nielsen, J. M., L. A. Rogers, R. D. Brodeur, A. R. Thompson, T. D. Auth, A. L. Deary, J. T. Duffy-Anderson, M. Galbraith, J. A. Koslow, and R. I. Perry. 2021. Responses of ichthyoplankton assemblages to the recent marine heatwave and previous climate fluctuations in several Northeast Pacific marine ecosystems. Global Change Biology 27(3):506–520.

R Core Team. 2023. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ren, X., W. Liu, A. Capotondi, D. J. Amaya, and N. J. Holbrook. 2023. The Pacific Decadal Oscillation modulated marine heatwaves in the Northeast Pacific during past decades. Communications Earth & Environment 4(1):1–9.

Rogers, L. A., and A. B. Dougherty. 2019. Effects of climate and demography on reproductive phenology of a harvested marine fish population. Global Change Biology 25(2):708–720.

Rogers, L. A., M. T. Wilson, J. T. Duffy-Anderson, D. G. Kimmel, and J. F. Lamb. 2021. Pollock and "the Blob": Impacts of a marine heatwave on walleye pollock early life stages. Fisheries Oceanography 30(2):142–158.

Rogers, L. A., and K. Axler. 2023. Larval Fish Abundance in the Gulf of Alaska 1981–2023. *In* Ferriss, B.E. 2023. Ecosystem Status Report 2023: Gulf of Alaska, Stock Assessment and Fishery Evaluation Report, North Pacific Fishery Management Council, 1007 West Third, Suite 400, Anchorage, AK 99501.

Simpson, G. 2023. gratia: Graceful ggplot-Based Graphics and Other Functions for GAMs Fitted using mgcv. R package version 0.8.1, https://gavinsimpson.github.io/gratia/.

Shotwell, S. K., M. Dorn, A. L. Deary, B. Fissel, L. Rogers, S. Zador, G. Adams, M. Arimitsu, K. Aydin, S. Barbeaux, L. Barnett, D. Hanselman, K. Holsman, D. Kimmel, B. Laurel, J. Pirtle, P. Ressler, R. Suryan, J. Thorson, J. Vollenweider, and S. Wise. 2019. Ecosystem and socioeconomic profile of the walleye pollock stock in the Gulf of Alaska, p. 104-150. *In* Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska. North Pacific Fishery Management Council, 605 W 4th Ave, Suite 306 Anchorage, AK 99501.

Suryan, R. M., M. L. Arimitsu, H. A. Coletti, R. R. Hopcroft, M. R. Lindeberg, S. J. Barbeaux, S. D. Batten, W. J. Burt, M. A. Bishop, J. L. Bodkin, R. Brenner, R. W. Campbell, D. A. Cushing, S. L. Danielson, M. W. Dorn, B. Drummond, D. Esler, T. Gelatt, D. H. Hanselman, S. A. Hatch, S. Haught, K. Holderied, K. Iken, D. B. Irons, A. B. Kettle, D. G. Kimmel, B. Konar, K. J. Kuletz, B. J. Laurel, J. M. Maniscalco, C. Matkin, C. A. E. McKinstry, D. H. Monson, J. R. Moran, D. Olsen, W. A. Palsson, W. S. Pegau, J. F. Piatt, L. A. Rogers, N. A. Rojek, A. Schaefer, I. B. Spies, J. M. Straley, S. L. Strom, K. L. Sweeney, M. Szymkowiak, B. P. Weitzman, E. M. Yasumiishi, and S. G. Zador. 2021. Ecosystem response persists after a prolonged marine heatwave. Scientific Reports 11(1):6235.

Thorson, J. T., C. J. Cunningham, E. Jorgensen, A. Havron, P.-J. F. Hulson, C. C. Monnahan, and P. von Szalay. 2021. The surprising sensitivity of index scale to delta-model assumptions: Recommendations for model-based index standardization. Fisheries Research 233:105745.

Wood, S. N. 2017. Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC.

Zador, S. G., K. K. Holsman, K. Y. Aydin, and S. K. Gaichas. 2017. Ecosystem considerations in Alaska: the value of qualitative assessments. ICES Journal of Marine Science 74(1):421–430.

Zhen, Z., L. Shao, and L. Zhang. 2018. Spatial hurdle models for predicting the number of children with lead poisoning. International Journal of Environmental Research and Public Health 15(9):1792.

U.S. Secretary of Commerce
Gina M. Raimondo

Under Secretary of Commerce
for Oceans and Atmosphere
Dr. Richard W. Spinrad

Assistant Administrator, National
Marine Fisheries Service.
Janet Coit

May 2024

www.nmfs.noaa.gov

OFFICIAL BUSINESS

National Marine
Fisheries Service
Alaska Fisheries Science Center
7600 Sand Point Way N.E.
Seattle, WA 98115-6349