Predicting year class strength for climate-stressed gadid stocks in the Gulf of Alaska

Michael A. Litzow[a]*, Alisa A. Abookire[b], Janet Duffy-Anderson[c], Benjamin J. Laurel[d], Michael J. Malick[e], Lauren A. Rogers[f]

[a]National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, 301 Research Ct. Kodiak, AK 99615, USA; mike.litzow@noaa.gov

[b]Alaska Coastal Observations and Research, PO Box 1332, Kodiak, AK 99615, USA; alaskacor@gmail.com

[c]National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115, USA; janet.duffy-anderson@noaa.gov

[d]National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, Hatfield Marine Science Center, 2030 SE Marine Science Dr., Newport, OR 97365, USA; ben.laurel@noaa.gov

[e]National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northwest Fisheries Science Center, 7305 Beach Dr. East, Port Orchard, WA 98366, USA; michael.malick@noaa.gov

[f]National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115, USA; lauren.rogers@noaa.gov

*Corresponding author

Abstract

Climate change makes fish stocks more vulnerable to recruitment failure, and early detection of these events is important for an effective management response. Here, we evaluate the value of larval and juvenile surveys, and a thermal spawning habitat index, for predicting recruitment in two economically important gadids, walleye pollock (*Gadus chalcogrammus*) and Pacific cod (*G. macrocephalus*), in the Gulf of Alaska. These stocks have been exposed to rapid human-induced ocean warming since 2014, which has apparently contributed to anomalies in age structure, size at age, and other population variables (for pollock) and stock collapse (for cod). We found that warming results in recruitment that falls short of predictions from historical spawner-recruit relationships for both stocks, highlighting that climate change makes recruitment expectations based on historical experience less reliable. However, we also found that recruitment could be successfully predicted with surveys of early-life stages. Using Bayesian regression, we found that juvenile trawl survey data for pollock predicts recruitment to age-1 (as estimated by a stock assessment model), while prediction from larval surveys was less successful. Beach seine estimates of juvenile abundance also predicted pollock recruitment, a surprising result for a species that is typically sampled in offshore habitats. The spawning habitat index and beach seine survey both predicted cod recruitment to age-3 as estimated by the stock assessment model. We did not find a predictive relationship between cod larval abundance and recruitment. However, residuals from the larval model showed low-frequency variability, suggesting nonstationarity (time-dependence) in the predictive relationship. Dynamic Factor Analysis (DFA) models summarizing information across multiple data sets showed reasonable predictive value for both species (Bayesian $R^2 \approx 0.4$ for log recruitment), and they also allowed recruitment prediction for years with missing observations in some data sets. We conclude that surveying multiple early life stages may be the most useful approach for predicting gadid recruitment.

Keywords: climate change; early life stages; Pacific cod; recruitment; walleye pollock

1. Introduction

The signal of human-induced climate change is emerging from the envelope of natural variability for many ocean ecosystems (Henson et al. 2017, Laufkötter et al. 2020, Silvy et al. 2020), placing fish stocks around the world under unprecedented climate conditions. Expected outcomes include an increase in the rate of events such as stock collapse (Pershing et al. 2019). Critically, these climate change outcomes may be “surprising” in that they are poorly constrained by ecological understanding based on historical experience. In this context, “historical” experience is derived from a backward-looking perspective on ecosystem and population variability, based on previous observations, which contrasts with a forward-looking perspective that assumes that current trends will continue (Dietze et al. 2018, Pershing et al. 2019). As anthropogenic climate extremes place ecosystems around the world under conditions that are outside the envelope of historical variability, ecological understanding based on historical conditions is increasingly less relevant, and tools are needed for rapidly assessing ecosystem states that have never been observed before (Williams and Jackson 2007, Wolkovich et al. 2014, Dietze et al. 2018). One important tool for avoiding collapse in climate-stressed stocks is early recognition of changing population dynamics to allow rapid adaptation by managers (Pershing et al. 2015). Changes to recruitment (the production of young fish to sustain the stock) are particularly important for early detection of incipient population changes. Fishing truncates the age structure of exploited stocks (Barnett et al. 2017), which has the pernicious effect of increasing recruitment variance and sensitivity to climate perturbations while also making stock biomass more dependent on regular recruitment events (Anderson et al. 2008, Shelton and Mangel 2011). Predicting recruitment (i.e., year class strength) is therefore of particular interest for the management of climate-stressed fisheries.

Recruitment and other population parameters are often estimated with age-structured stock assessment models, which are advantageous for combining information from different data sources. However, out-of-sample prediction of recruitment by these models is often very low (Giron-Nava et al. 2020). Recruitment estimates are therefore typically made retrospectively, after cohorts have appeared in the fishery or fishery-independent surveys. A lag therefore exists between the timing when year class strength is established, which typically occurs at the first year of life (age-0; Houde 2008), and recognition of year class strength by the model used for estimating management reference points. Observations of cohorts at age-0 (as eggs, larvae, or juveniles) might therefore be useful for providing earlier indications of future adult recruitment. However, extreme levels of sampling variance and highly variable early life mortality have often precluded effective prediction of recruitment by early life stage surveys (Stige et al. 2013).

Here, we evaluate the ability of larval and juvenile abundance surveys and a temperature-based index of spawning habitat suitability to predict recruitment as estimated by stock assessment models for two exploited gadids in the Gulf of Alaska: walleye pollock (*Gadus chalcogrammus*, hereafter “pollock”) and Pacific cod (*G. macrocephalus*, hereafter “cod”). These populations have been the subject of early life stage monitoring efforts for forty years and so provide a data-rich model system for examining this question. Early detection of recruitment variability for these stocks has become more important since a series of extreme ocean temperature anomalies in the Gulf of Alaska during 2014-2019, which have been formally attributed to human-induced causes (Walsh et al. 2018b, Laufkötter et al. 2020). The transition from conditions consistent with natural variability to extreme ocean temperatures that can only be explained with human influence (e.g., greenhouse gas emissions) is associated with an increased likelihood of recruitment failure for both species (Litzow et al. 2021). This anthropogenic warming event apparently contributed to a number of anomalous responses by the pollock stock, including negative weight-at-age anomalies, early maturation and reduced natural mortality, and strongly contrasting abundance trends in bottom trawl and acoustic midwater surveys (Dorn et al. 2020). The warming event also resulted in collapse of the cod stock, apparently due to the impacts of early life stage thermal stress combined with increased metabolic demands and insufficient prey resources for older cohorts (Barbeaux et al. 2020b), though synergistic fishing effects cannot be ruled out (Harley et al. 2006, Hsieh et al. 2006).

There is a long history of research on the relationship between early life stage abundance and recruitment to the pollock stock. In particular, Bailey et al. (2012) and Stige et al. (2013) showed that estimates of larval abundance alone were poor predictors of subsequent recruitment. However, these authors also found that the predictive value for these time series was nonstationary (i.e., time-varying). This nonstationarity appears to result from changing control of pollock recruitment over time. In the 1980s, larval mortality was the primary driver of recruitment variability (Bailey 2000), but in the 1990s predation on juveniles became a more important limiting factor for recruitment, and the predictive value of larval abundance degraded (Ciannelli et al. 2004, 2005). More recently, abundance of age-0 pollock has been shown to be an early indicator of year class strength in this stock (Wilson and Laman 2021).

Less work has focused on the predictive value of early life stage observations for cod recruitment. Laurel et al. (2016) showed that beach seine catches of age-0 cod could predict age-1 abundance at the same sites in the following year. Laurel and Rogers (2020) showed that an index of spawning habitat suitability (depth-averaged temperature conditioned on laboratory estimates of temperature-dependent egg survival) could predict larval cod abundance, juvenile cod abundance from seines at two long-term sampling sites at Kodiak Island, and recruitment estimates from the stock assessment model. This analysis did not consider the value of larval abundance or juvenile abundance from seines for predicting modeled recruitment, and it also did not consider an additional set of seines over a wider area of the western Gulf of Alaska (Litzow et al. 2021). Additionally, post-2014 climate anomalies in the Gulf of Alaska may have sharply reduced habitat suitability for larval cod due to changes in timing of the spring bloom and increased metabolic demands of larvae, which may have further disrupted predictive relationships between larval abundance and cod recruitment (Laurel et al. 2021).

Thus while strong climate forcing on these stocks increases the need for better recruitment prediction, the magnitude of climate perturbations also increases the possibility of changing control points for cohort strength, increasing the chance of nonstationary predictions from early life stage observations. Re-evaluation of the predictive value of early life stage observations is therefore needed to establish whether and how these data can be used quantitatively in stock assessments or management strategy evaluations, or qualitatively to inform management advice (Dorn and Zador 2020).

The goal of the current study is to re-assess the value of early life stage observations for predicting pollock and cod recruitment. Our analysis puts post-2014 climate extremes in the context of four decades of observations, and it also draws on recent advances in dimensional reduction techniques for time series data to combine information from multiple data sources for individual stocks. Our specific objectives are to: 1) use Bayesian models to produce probabilistic assessments of the value of individual early life stage time series for predicting recruitment estimates from stock assessment models, 2) assess predictive value for Dynamic Factor Analysis (DFA) models that combine information across multiple time series, and 3) use time series of model residuals to explore possible nonstationarity in predictions. We also examine the effect that extreme climate anomalies have on predictions generated from historical stock-recruit relationships, as a way of assessing the vulnerability of expectations based on historical norms during unprecedented climate conditions (Pershing et al. 2019).

2. Methods

2.1. Data sources

We used four sources of biological information: an index of spawning habitat suitability for cod based on temperature-dependent egg survival at depth (Laurel and Rogers 2020), a May-June ichthyoplankton survey that samples larvae of both species, July-August beach seines that sample juveniles of both species at fixed site locations, and an August-September pelagic trawl survey that samples juvenile pollock.

The index of spawning habitat suitability for cod was constructed by combining observational data on temperatures at depth with laboratory-derived rates of hatch success of cod eggs at different temperatures (see Laurel and Rogers, 2020). Specifically, measured temperatures at GAK1, a long-term monitoring station maintained by the University of Alaska (http://research.cfos.uaf.edu/gak1/; Fig. 1a), were used to estimate hatch success across depths, seasons, and years. An annual index was developed by taking the average estimated hatch success rate across depths of 100–250 m from January to April based on reported spawning dynamics for cod in Alaska (Stark 2007, Neidetcher et al. 2014).

Fish larvae were sampled during 1981-2019 (missing years in 1984, 1986, 2012, 2014, 2016, and 2018), and juvenile pollock trawl surveys were conducted during 2000-2001 and odd years from 2003 to 2019. Larvae were sampled over a fixed area (Fig. 1a) during mid-May — early June using oblique tows from 10 m off bottom (or 100 m depth maximum) to the surface using a 60-cm diameter bongo net (333 or 505 µm mesh). Calibrated flowmeters in each net estimated the volume filtered, and catch was standardized as number 10 m^{-2} area of sea surface sampled. Finally, time series indices of larval abundance were calculated as the area-weighted mean catch to account for spatial differences in sampling effort among years (Doyle et al. 2009).

Beach seines were conducted with a negatively buoyant, 36-m long seine, with wings 1 m deep at the ends of the net and 2.25 m deep in the middle, 13 mm mesh in the wings and 5 mm delta mesh in the cod end bag. Seine wings were attached to 25 m ropes for deployment and retrieval from shore, and nominal sampling area was ≈ 900 m^2 of bottom habitat. Sampling was conducted during July and August at 95 sites in 15 bays (Fig. 1a), with each site sampled 1-4 times per year. The two easternmost bays were sampled each year during 2006-2020, with at least two sampling visits per year (n = 880 sets). The remainder of bays were sampled during 2018-2020 (n = 265 sets). Because beach seine effort varied spatially and seasonally in different years, we used model-based estimates of annual abundance for each species as our time series for subsequent analyses. These estimates were the predicted annual catch per unit effort (fish / set) from Bayesian regression models invoking year as a factor, day of year as a smooth term (thin plate regression splines) and site nested within bay (as a group-level, or random term). In particular, we used a Generalized Additive Model (GAM) that estimates annual abundance after controlling for the day of year of each sampling event (as a smoothed non-parametric term to allow for nonlinear seasonal changes in abundance), and nested site and bay random effects (to account for differences in abundance among sites and bays that are sampled unequally across years). The Bayesian formulation has the benefit of providing full uncertainty estimates of model coefficients and predicted values (Fang et al. 2019). Details on Bayesian model fitting are given in section 2.2, and further details are in Litzow et al. (2021).

Trawl surveys for age-0 pollock were conducted in a fixed area (Fig. 1a) in August-September. Samples were collected using a midwater trawl fished with 1.5 × 2.1 m steel V-doors (566 kg each) and equipped with a 3 mm cod end liner. The trawl was fished obliquely through the water column at a ship speed of 4.6 to 5.6 km hr^{-1} and a wire retrieval rate of 10 m min^{-1}. A time series of age-0 pollock abundance was developed by calculating an area-weighted mean catch m^{-2} in each year, using the same methodology as for the larval index (Rogers et al. 2021).

Stock assessment model estimates of recruitment at age-0 are from the authors' preferred models from 2020 stock assessment reports. These stock assessment models are age-structured models that estimate a number of population parameters using a variety of data inputs; models are fit in Stock Synthesis and AD Model Builder software and form the basis for setting management reference points and fishing quotas (details in Barbeaux et al. 2020a, Dorn et al. 2020). Recruitment estimates are lagged from estimated age-1 abundance for pollock, and age-3 abundance for cod. This difference in the age at which recruitment is estimated reflects differences in data availability for the two models: age-1 pollock are sampled by acoustic trawl surveys, but the earliest life-stage data for the cod model are provided when cod begin to appear in bottom trawl sampling gear at age-3. Our analysis only included recruitment estimates that overlapped with observational time series and were supported by data on year class strength within the model (year classes 1981-2019 for pollock, 1981-2016 for cod).

Our historical SST data come from the NOAA Extended Reconstructed SST data set, version 5 (ERSSTv5; Huang et al. 2017). We calculated annual mean SST values for the western Gulf of Alaska (the area of our pollock and cod sampling) for the months of January – June (corresponding the spawning, larval, and early juvenile phases for the two species). Projected future SST anomalies were calculated for the Gulf of Alaska from downscaled climate model outputs for five CMIP5 models (Climate Model Intercomparison Project, Phase 5) with good predictive skill for Alaska (Walsh et al. 2018b, 2018a). These anomalies are from combined time series of model hindcasts (1987-2005) and projections under Representative Concentration Pathway (RCP) 8.5 (Schwalm et al. 2020).

2.2. Analysis

We began our analysis by evaluating the degree to which extreme climate anomalies resulted in error for recruitment predictions based on historical norms. We simulated recruitment predictions for each stock on a rolling window basis using the following approach. We fit a Ricker model to the first 30 years of the estimated recruitment and spawning stock biomass time series, and we then used that Ricker model to predict recruitment in the following year (i.e., recruitment predicted from spawning stock biomass in year 31). We also calculated the magnitude of the SST anomaly in the prediction year, based on the mean and standard deviation of SST values for the same 30 year period. The recruitment prediction error was recorded (i.e., ln[predicted recruitment for year 31] - ln[actual modeled recruitment estimate for year 31]), as was the SST anomaly. The window was then advanced one year (i.e., recruitment given spawning stock biomass in year 32 was predicted from a Ricker model fit to the first 31 years of data, and the SST anomaly was calculated in year 32 relative to the first 31 years), and this procedure was repeated until the end of the stock assessment model time series was reached. Note that our goal in this analysis was to evaluate the impact of warming on predictions that might be made from historical spawner-recruit relationships, rather than assessing the ability of stock assessment models to predict out-of-sample recruitment (Giron-Nava et al. 2020). We used the rolling window approach because it explicitly models a situation where historical understanding (based on data from the rolling window) is confronted with new information (from the update year). Recruitment errors were scaled as z-scores for each species (difference from the mean value, divided by the standard deviation), and the relationship between SST anomalies and recruitment prediction error, for both species combined, was estimated with Bayesian regression, using thin plate splines to model a non-parametric relationship (Wood 2003). An initial model that fit the SST – recruitment prediction error relationship separately to each species showed that there was no substantive difference in effects between species. This model with species-specific effects produced worse out-of-sample prediction (i.e., a higher Leave One Out Information Criterion score; Vehtari et al. 2017), and was therefore rejected in favor of a model that pooled results across species.

After the relationship between SST anomalies and recruitment prediction error was estimated, we summarized hindcasted and projected probabilities of extreme SST anomalies (> 2SD and > 3SD) for annual Gulf of Alaska SST during 1987-2046. This analysis used a combination of historical simulations (1987-2005) and RCP8.5 projections (2006-2046) from CMIP5 outputs. For each year, we estimated the probability of extreme events as the proportion of the five models that projected anomalies >2 SD and >

3SD (Walsh et al. 2018b). The resulting time series were smoothed with 11-year rolling means to isolate the trend from the noise of interannual variability.

All time series of observed abundance and modeled recruitment strength were normalized with natural log transformations prior to analysis. All time series (including the habitat index) were scaled as z-scores to aid comparison across data types. Normalized, scaled data are plotted for each species in Fig. 1b,c.

We used Dynamic Factor Analysis (DFA; Zuur et al. 2003) to summarize variability across the observational time series for each species (larval surveys, seines, and trawl surveys for pollock; habitat index, larval surveys, and seines for cod). DFA is a dimensional reduction technique developed specifically for time series analysis which estimates variability in an unobserved (or “latent”) shared trend, based on loadings for individual time series and a variance-covariance matrix. We fit DFA models in the MARSS package version 3.11.3 (Holmes et al. 2012) in R version 4.0.4 (R Core Team 2021). We fit two different variance-covariance structures for each species (same variance and no covariance or different variance and no covariance), with the best model selected based on the Akaike Information Criterion. Two other model structures (same variances and same covariance, different variances and covariances) returned models with loadings of 0 and were dropped from consideration. DFA models were fit to that part of the time series that included at least one observation in every year (1987-2020 for pollock, 1994-2020 for cod).

When evaluating the ability of different early-life data sets to predict recruitment, we no longer used the rolling window approach but simply used all of the data in hand. This is because we were no longer interested in challenges to historical understanding, but instead were interested in the ability of early-life data to predict eventual recruitment. We estimated predictive value for each observational time series, and for the shared trend from DFA models, through comparison with stock assessment model recruitment estimates using Bayesian linear regression models fit in Stan 2.21.0 and the brms package in R (Carpenter et al. 2017, Bürkner 2017). All estimated parameters had a potential scale reduction factor ($\hat{R}$) less than 1.05, an effective sample size of at least 1000, and no divergent transitions were observed. We also assessed chain convergence and model fits using graphical methods (e.g., trace-plots) and posterior predictive checks (Gabry et al. 2019). Different years were covered by each time series (Fig. 1b,c), so formal model comparison (e.g., by the Leave One Out Information Criterion) was not feasible (i.e., different response variable sets were available for each year). Rather than model comparison, we evaluate the predictive value of different time series for stock assessment model recruitment estimates by comparing regression coefficients and Bayes R^2 values, which are calculated from the model posteriors as

$$R^2 = \frac{Var_{\mu}}{Var_{\mu} + Var_{res}},$$

where Var_{μ} is the variance of modelled predictive means, and Var_{res} is the modelled residual variance (Gelman et al. 2019). While we recognize that these metrics do not support as rigorous a comparison as formal model selection techniques, we judged that this approach was superior as it used all of the information available for each time series.

Finally, we conducted an exploratory analysis of potential nonstationarity in predictive value for models fit to data from each observational stage (i.e., spawning habitat suitability, larval and juvenile abundance) by plotting residual time series from model posteriors (with 95% CIs). Temporal trends in model residuals are an important indication that a model assuming stationary (time-independent) relationships is inadequate to a situation where ecological relationships (in this case, the relationship between early-life data and eventual recruitment) are changing over time. Conversely, independently distributed residuals (without temporal dependence) are an indication that the assumption of stationarity is valid (Litzow et al. 2018, Rollinson et al. 2021). To examine potential time-dependent errors in predictions we fit GAMs and associated confidence intervals (assuming independence among years) to the time series of residual means. Periods when the confidence interval for the GAM fits did not include 0 were judged to reflect possible changes in prediction value. All data and code necessary for reproducing our results are posted in the "predict-R" repository (https://github.com/mikelitzow/predict-R) [Permanent repository in Zenodo on acceptance].

3. Results

Strong climate change events (as indexed by SST anomalies) were associated with failure of recruitment predictions generated from historical spawner-recruit relationships (Bayes R^2 = 0.42, 95% CI = 0.16 - 0.58). SST anomalies > 2SD were associated with the largest prediction errors (actual recruitment more than 1 SD less than predicted; Fig. 2a). Downscaled CMIP5 models indicate that the probability of these extreme temperatures associated with unpleasant "surprises" (low recruitment events poorly constrained by historical spawner-recruit relationships) is rapidly increasing (e.g., probability of anomalies > 2SD ≈ 0.4 by 2030, ≈ 0.6 by 2040). The annual probability of more extreme anomalies (> 3SD) is projected to reach ≈ 0.2 by 2040 (Fig. 2b).

For pollock, the larval abundance, age-0 seine abundance, and age-0 trawl abundance time series were all informative for estimating the shared trend in recruitment variability, as indicated by positive DFA loadings with 95% CI that did not include 0 (Fig. 3a). The DFA-estimated trend in pollock recruitment strength showed a run of negative values during the 2014-2016 heatwave years (Fig. 3b). For cod, all three time series also showed positive loadings that could be distinguished from 0 in the DFA model, with some indication of stronger loadings for the earliest time series (habitat index and larval abundance) and a weaker loading for the later time series (age-0 seine; Fig. 3c). The shared trend in the cod DFA model indicated a step decline in recruitment since the onset of temperature extremes in 2014 (Fig. 3d).

Bayesian regression coefficients and Bayes R^2 values indicated a gradient of predictive value in pollock time series, with weaker predictions based on larval abundance, intermediate predictions for age-0 seines, and very strong predictive value for age-0 trawls (Fig. 4). The DFA trend for pollock showed predictive value intermediate between that of larval abundance and age-0 trawls, and regression coefficients for all four time series could be distinguished from 0 (Fig. 4). Residual time series for the four pollock models showed a declining trend, but this trend could be distinguished from 0 only for the larval time series (Fig. 5). The final four years of the time series (2016-2019) included the three lowest residuals from the larval model, and the magnitude of these residuals was much greater than that of persistently positive (negative) residuals during the 1980s (1990s).

Beach seines showed the strongest predictive value for cod recruitment, with markedly weaker predictive value for the habitat index, and a regression coefficient for larval abundance that could not be distinguished from 0 (Fig. 6). The DFA trend for cod was intermediate in predictive value. Residual time series for the larval prediction model showed a transition from persistently positive values in the 1980s to persistently negative values in the 1990s through the mid-2000s, followed by residuals that were not persistently negative or positive during 2007-2016 (Fig. 7). Residuals from the cod DFA model showed some evidence of a transition from negative to positive values across the 1994-2016 time series (Fig. 7).

4. Discussion

All three early life stage time series showed value for predicting recruitment to the pollock stock, and two of three time series showed value for predicting recruitment to the cod stock, as judged by 95% credible intervals for regression coefficient estimates that did not include 0. The exception was prediction of cod recruitment from the larval survey, and this was also the time series that showed the strongest indication of a nonstationary relationship with recruitment. This result indicates that the relationship between larval abundance and cod recruitment has changed over time. A similar trend from negative to positive residuals was observed for predictions from the spawning habitat suitability index, indicating coherence in this result for early-life cod (i.e., similar nonstationary trends in prediction residuals for both time series). And, since the habitat and larval time series load heavily on the DFA trend for cod (Fig. 3), a similar trend in prediction residuals was observed for the DFA trend (Fig. 7). Accounting for this changing relationship might provide an avenue for improving recruitment prediction, and this nonstationarity also suggests that research on changing controls of recruitment may improve inference about evolving controls of population dynamics for this stock (Rollinson et al. 2021). We also found some indication of trends towards negative residuals for pollock prediction models, particularly for the larval time series, which may indicate a trend of worsening survival between early life stages and recruitment to that stock.

We conclude that early life observations were generally useful for predicting recruitment in our study populations, though nonstationarity in predictive relationships should be regularly evaluated. We found that errors for predictions made without early life stage data (i.e., from historical spawner-recruit relationships) are strongly associated with the magnitude of climate anomalies acting on the system (Fig. 2a), and that the probability of strong climate anomalies is rapidly increasing (Fig. 2b). Given these trends, historical expectations concerning the envelope of likely variability for these stocks are becoming rapidly outmoded (Pershing et al. 2019, Litzow et al. 2021). These considerations make early information on recruitment variability especially valuable (Pershing et al. 2015). In particular, observations of very low abundance at the larval or juvenile stages appear to be a reliable indicator of a weak year class. Examples for pollock include 2015 and 2019 (Fig. 4) and 2015 and 2016 for cod (Fig. 6). The extreme climate conditions during these years were associated with novel conditions affecting year-class strength – both in abiotic conditions (low salinity reducing pollock egg survival, high temperatures reducing cod egg survival), and biotic conditions (Laurel and Rogers 2020, Laurel et al. 2021, Rogers et al. 2021).

The practical value of predictions from early life stage observations was evident during the recent marine heatwaves, when these data were used to provide early warning of likely recruitment failures for these stocks, 1-3 years before such impacts were detectable in fisheries or standard survey data used in stock assessments. While early life stage data are not routinely included in stock assessment models in the Gulf of Alaska, these data are increasingly used to provide supplemental information for the stock assessment teams giving management advice, and for the managing body, the North Pacific Fisheries Management Council (NPFMC), which makes decisions on annual fishing quotas. In 2019, the cod stock assessment lead identified concerns that continued lack of recruitment would delay recovery of the stock following the steep decline in 2015-2017. The management system for Pacific cod ensures that catch limits decrease as the population declines based on a harvest control rule. However, the NPFMC took extra precaution and reduced catch levels an additional 40%, citing concerns about a weak 2019 year class and the subsequent risk that the stock might fall into "overfished" status in 2020 or 2021. By providing rapid assessments of potential year class strength, early life stage surveys can thus be used to guide tactical fisheries management decisions, particularly during climate events that may be outside the range of historical variability.

As judged by coefficient estimates, the DFA trend for pollock showed intermediate predictive value for pollock (slightly better than larval and juvenile seine prediction, worse than juvenile trawl prediction; Fig. 4). Larval and juvenile surveys are only conducted in odd years, so updated values of the DFA trend, informed by the annual seine data, may be useful for predicting recruitment for even-year cohorts. In contrast, juvenile seine data were superior to the spawning habitat index and larval survey for predicting cod recruitment (Fig. 6), so there is no advantage to be gained from using the DFA trend over seine data in that case.

Cod and pollock population trends in Alaska tend to be correlated over interannual to decadal time scales (Hollowed et al. 2001, Mueter et al. 2007), and that tendency was born out in our data - the pollock DFA trend and cod log seine abundance were correlated at $r = 0.73$. However, while cod and pollock are morphologically similar at early life stages, they show distinctive behavior and physiology that will likely elicit different responses to shared climate forcing. Pollock release eggs over the course of a spring season in the form of multiple batches from individual females, and pollock eggs are relatively ubiquitous in surface layers during ichthyoplankton spring surveys (Doyle and Mier 2016, Rogers et al. 2021). In contrast, cod are single-batch spawners with demersal eggs that are deposited on the ocean floor. Cod year class strength is less likely to be impacted by variability in advection fields and surface temperature following spawning, but cod eggs are much more sensitive to warm anomalies at depth (Laurel and Rogers 2020). After hatch, both Pacific cod and walleye pollock share the water column as larvae (Doyle and Mier 2016) and have similar growth response to temperature (Hurst et al. 2010, Laurel et al. 2016, Koenker et al. 2018). Following the larval period, the life histories of these species diverge again as Pacific cod go through a transition to life on the ocean floor and begin a period of nearshore occupancy ('settlement') while juvenile pollock remain in the surface layers in both nearshore and offshore habitats (Moss et al. 2015; Laurel et al. 2016).

As warming and other anthropogenic climate change effects move the Gulf of Alaska ecosystem further from the historical range of conditions, these differences in life history, habitat occupation and physiology suggest the potential for species-specific patterns of further nonstationarity in the predictive relationships demonstrated in this study. Relationships among temperature, advection, and salinity have changed historically in the Gulf of Alaska in ways that have apparently contributed to nonstationary climate-biology relationships across the fish and crustacean community (Litzow et al. 2018, 2019), and further changes to these climate relationships may have accompanied the extreme 2014-2019 temperature anomalies (Litzow et al. 2020). Previous nonstationarity in recruitment prediction from pollock larval abundance (Bailey 2000, Stige et al. 2013) is apparent in our data in a run of positive residuals in the 1980s followed by more neutral residuals in the 1990s (Fig. 5). However, the magnitude of those earlier residual trends appears to be much less than the magnitude of negative residuals during 2014-2019. These recent negative residuals for recruitment predictions from pollock larval abundance appear to signal declining early life stage survival in the contemporary Gulf of Alaska climate (Rogers et al. 2021). Residuals for cod recruitment predictions from larval abundance also showed a previously-unreported low-frequency pattern of persistently positive values in the 1980s and negative values in the 1990s and 2000s (Fig. 7). The timing of the change in sign of these residuals is consistent with both a decline in temporal variance in the Aleutian Low, which appeared to have widespread implications for nonstationary regulation of Gulf of Alaska fish populations (Litzow et al. 2018, 2019), and also with the proposed timing of the switch in the primary factor regulating pollock recruitment (Bailey 2000, Stige et al. 2013). The coincident change in cod residuals might imply that that species also saw a transition from recruitment limited by larval mortality to recruitment limited by predation on juveniles. Such a transition would be consistent with the proposed transition for control of pollock recruitment around the same time, from environmental control of larval abundance to predation control of juvenile abundance (Bailey 2000). However, we do not have data in hand to test that hypothesis in the current study. Measurement error may also be an important source of nonstationarity in predictive relationships. Climate change effects phenology and growth of early life stage fishes (Laurel et al. 2021), which may result in an increasing disconnect between the timing of fixed-design surveys and the peak abundance of different life history stages they target.

If projections concerning the expected rate of warming in the North Pacific and Gulf of Alaska are correct (Walsh et al. 2018b, Laufkötter et al. 2020), and given that these populations are at the southern limit of their range as commercially-important fisheries, the most likely prediction for both species in coming decades appears to be persistent recruitment failure and local extirpation, at least commercially. An important aspect of successful climate adaptation by fisheries stakeholders is the maintenance of catches from declining stocks for as long as possible to give time for new fisheries to be developed (Cinner et al. 2018). Predicting recruitment to these stocks using early-life surveys may be particularly valuable in this context. Given the potential for nonstationary predictive value for individual time series, monitoring across multiple early life stages will likely be the most effective approach for developing robust predictions of year class strength as climate change effects on these stocks intensify.

CRediT authorship contribution statement

Michael A. Litzow: Conceptualization, Formal analysis, Investigation, Software, Visualization, Writing – original draft, Writing – review & editing. **Alisa A. Abookire:** Investigation, Writing – review & editing. **Janet Duffy-Anderson:** Investigation, Writing – review & editing. **Benjamin J. Laurel:** Investigation, Writing – review & editing. **Michael J. Malick:** Investigation, Software, Validation, Visualization, Writing – review & editing. **Lauren A. Rogers:** Investigation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the many people who helped with the decades of fieldwork necessary for collecting these data, and we thank John Walsh for sharing CMIP5 output. Funding was provided by the NOAA Alaska Fisheries Science Center and a NOAA Cooperative Research Grant. The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of the National Marine Fisheries Service.

References

Anderson, C. N. K., C. H. Hsieh, S. A. Sandin, R. Hewitt, A. Hollowed, J. Beddington, R. M. May, and G. Sugihara. 2008. Why fishing magnifies fluctuations in fish abundance. Nature 452:835–839.

Bailey, K. M. 2000. Shifting control of recruitment of walleye pollock Theragra chalcogramma after a major climatic and ecosystem change. Marine Ecology Progress Series 198:215–224.

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

Barbeaux, S. J., B. Ferriss, W. Palsson, K. Shotwell, I. Spies, M. Wang, and S. Zador. 2020a. Assessment of the Pacific cod stock in the Gulf of Alaska.

Barbeaux, S. J., K. Holsman, and S. Zador. 2020b. Marine heatwave stress test of ecosystem-based fisheries management in the Gulf of Alaska Pacific cod fishery. Frontiers in Marine Science 7:703.

Barnett, L. A. K., T. A. Branch, R. A. Ranasinghe, and T. E. Essington. 2017. Old-growth fishes become scarce under fishing. Current Biology 27:2843–2848.

Bürkner, P. C. 2017. brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software 80:1–28.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. A. Brubaker, J. Guo, P. Li, and A. Riddell. 2017. Stan: A Probabilistic Programming Language. Journal of Statistical Software 76:1–29.

Ciannelli, L., K. M. Bailey, K. S. Chan, A. Belgrano, and N. C. Stenseth. 2005. Climate change causing phase transitions of walleye pollock (Theragra chalcogramma) recruitment dynamics. Proceedings of the Royal Society B-Biological Sciences 272:1735–1743.

Ciannelli, L., K. S. Chan, K. M. Bailey, and N. C. Stenseth. 2004. Nonadditive effects of the environment on the survival of a large marine fish population. Ecology 85:3418–3427.

Cinner, J. E., W. N. Adger, E. H. Allison, M. L. Barnes, K. Brown, P. J. Cohen, S. Gelcich, C. C. Hicks, T. P. Hughes, J. Lau, N. A. Marshall, and T. H. Morrison. 2018. Building adaptive capacity to climate change in tropical coastal communities. Nature Climate Change 8:117–123.

Dietze, M. C., A. Fox, L. M. Beck-Johnson, J. L. Betancourt, M. B. Hooten, C. S. Jarnevich, T. H. Keitt, M. A. Kenney, C. M. Laney, L. G. Larsen, H. W. Loescher, C. K. Lunch, B. C. Pijanowski, J. T. Randerson, E. K. Read, A. T. Tredennick, R. Vargas, K. C. Weathers, and E. P. White. 2018. Iterative near-term ecological forecasting: Needs, opportunities, and challenges. Proceedings of the National Academy of Sciences of the United States of America 115:1424–1432.

Dorn, M. W., A. Deary, B. Fissel, D. T. Jones, M. Levine, A. L. McCarthy, and E. al. 2020. Assessment of the walleye pollock stock in the Gulf of Alaska.

Dorn, M. W., and S. G. Zador. 2020. A risk table to address concerns external to stock assessments when developing fisheries harvest recommendations. Ecosystem Health and Sustainability 6:1813634.

Doyle, M. J., and K. L. Mier. 2016. Early life history pelagic exposure profiles of selected commercially important fish species in the Gulf of Alaska. Deep-Sea Research Part II: Topical Studies in Oceanography 132:162–193.

Doyle, M. J., S. J. Picquelle, K. L. Mier, M. C. Spillane, and N. A. Bond. 2009. Larval fish abundance and physical forcing in the Gulf of Alaska, 1981-2003. Progress in Oceanography 80:163–187.

Fang, X., B. Fang, C. Wang, T. Xia, M. Bottai, F. Fang, and Y. Cao. 2019. Comparison of frequentist and Bayesian generalized additive models for assessing the association between daily exposure to fine particles and respiratory mortality: A simulation study. International Journal of Environmental

Research and Public Health 16:764.

Gabry, J., D. Simpson, A. Vehtari, M. Betancourt, and A. Gelman. 2019. Visualization in Bayesian workflow. Journal of the Royal Statistical Society Series A-Statistics in Society 182:389–402.

Gelman, A., B. Goodrich, J. Gabry, and A. Vehtari. 2019. R-squared for Bayesian regression models. American Statistician 73:307–309.

Giron-Nava, A., S. B. Munch, A. F. Johnson, E. Deyle, C. C. James, E. Saberski, G. M. Pao, O. Aburto-Oropeza, and G. Sugihara. 2020. Circularity in fisheries data weakens real world prediction. Scientific Reports 10:6977.

Harley, C. D. G., A. R. Hughes, K. M. Hultgren, B. G. Miner, C. J. B. Sorte, C. S. Thornber, L. F. Rodriguez, L. Tomanek, and S. L. Williams. 2006. The impacts of climate change in coastal marine systems. Ecology Letters 9:228–241.

Henson, S. A., C. Beaulieu, T. Ilyina, J. G. John, M. Long, R. Seferian, J. Tjiputra, and J. L. Sarmiento. 2017. Rapid emergence of climate change in environmental drivers of marine ecosystems. Nature Communications 8:14682.

Hollowed, A. B., S. R. Hare, and W. S. Wooster. 2001. Pacific Basin climate variability and patterns of Northeast Pacific marine fish production. Progress in Oceanography 49:257–282.

Holmes, E. E., E. J. Ward, and K. Wills. 2012. Marss: Multivariate autoregressive state-space models for analyzing time-series data. The R Journal 4:11–19.

Houde, E. 2008. Emerging from Hjort's Shadow. Fish. Sci. J. Northw. Atl. Fish. Sci 41:53–70.

Hsieh, C.-H., C. S. Reiss, J. R. Hunter, J. R. Beddington, R. M. May, and G. Sugihara. 2006. Fishing elevates variability in the abundance of exploited species. Nature 443:859–862.

Huang, B., P. W. Thorne, V. F. Banzon, T. Boyer, G. Chepurin, J. H. Lawrimore, M. J. Menne, T. M. Smith, R. S. Vose, and H.-M. Zhang. 2017. Extended reconstructed sea surface temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. Journal of Climate 30:8179–8205.

Hurst, T. P., B. J. Laurel, and L. Ciannelli. 2010. Ontogenetic patterns and temperature-dependent growth rates in early life stages of Pacific cod (Gadus macrocephalus). Fishery Bulletin 108:382–392.

Koenker, B. L., B. J. Laurel, L. A. Copeman, and L. Ciannelli. 2018. Effects of temperature and food availability on the survival and growth of larval Arctic cod (Boreogadus saida) and walleye pollock (Gadus chalcogrammus). ICES Journal of Marine Science 75:2386–2402.

Laufkötter, C., J. Zscheischler, and T. L. Frölicher. 2020. High-impact marine heatwaves attributable to human-induced global warming. Science 369:1621–1625.

Laurel, B. J., M. E. Hunsicker, L. Ciannelli, T. P. Hurst, J. Duffy-Anderson, R. O'Malley, and M. Behrenfeld. 2021. Regional warming exacerbates match/mismatch vulnerability for cod larvae in Alaska. Progress in Oceanography 193:102555.

Laurel, B. J., B. A. Knoth, and C. H. Ryer. 2016. Growth, mortality, and recruitment signals in age-0 gadids settling in coastal Gulf of Alaska. ICES Journal of Marine Science 73:2227–2237.

Laurel, B. J., and L. A. Rogers. 2020. Loss of spawning habitat and prerecruits of Pacific cod during a Gulf of Alaska heatwave. Canadian Journal of Fisheries and Aquatic Sciences 77:644–650.

Litzow, M. A., L. Ciannelli, P. Puerta, J. J. Wettstein, R. R. Rykaczewski, and M. Opiekun. 2018. Non-stationary climate-salmon relationships in the Gulf of Alaska. Proceedings of the Royal Society B: Biological Sciences 285:20181855.

Litzow, M. A., L. Ciannelli, P. Puerta, J. J. Wettstein, R. R. Rykaczewski, and M. Opiekun. 2019. Nonstationary environmental and community relationships in the North Pacific Ocean. Ecology 100.

Litzow, M. A., M. J. Malick, A. A. Abookire, J. Duffy-Anderson, B. J. Laurel, P. H. Ressler, and L. A. Rogers. 2021. Using a climate attribution statistic to inform judgments about changing fisheries sustainability. Scientific Reports 11:23924.

Litzow, M. A., M. J. Malick, N. A. Bond, C. J. Cunningham, J. L. Gosselin, and E. J. Ward. 2020. Quantifying a novel climate through changes in PDO-climate and PDO-salmon relationships. Geophysical Research Letters 47:e2020GL087972.

Mueter, F. J., J. L. Boldt, B. A. Megrey, and R. M. Peterman. 2007. Recruitment and survival of Northeast Pacific Ocean fish stocks: temporal trends, covariation, and regime shifts. Canadian Journal of Fisheries and Aquatic Sciences 64:911–927.

Neidetcher, S. K., T. P. Hurst, L. Ciannelli, and E. A. Logerwell. 2014. Spawning phenology and geography of Aleutian Islands and eastern Bering Sea Pacific cod (Gadus macrocephalus). Deep-Sea Research Part II: Topical Studies in Oceanography 109:204–214.

Pershing, A. J., M. A. Alexander, C. M. Hernandez, L. A. Kerr, A. Le Bris, K. E. Mills, J. A. Nye, N. R. Record, H. A. Scannell, J. D. Scott, G. D. Sherwood, and A. C. Thomass. 2015. Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery. Science 350:809–812.

Pershing, A. J., N. R. Record, B. S. Franklin, B. T. Kennedy, L. McClenachan, K. E. Mills, J. D. Scott, A. C. Thomas, and N. H. Wolff. 2019. Challenges to natural and human communities from surprising ocean temperatures. Proceedings of the National Academy of Sciences of the United States of America 116:18378–18383.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

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 30:142–158.

Rollinson, C. R., A. O. Finley, M. R. Alexander, S. Banerjee, K. A. Dixon Hamil, L. E. Koenig, D. H. Locke, M. Peterson, M. W. Tingley, K. Wheeler, C. Youngflesh, and E. F. Zipkin. 2021. Working across space and time: nonstationarity in ecological research and application. Frontiers in Ecology and the

Environment 19:66–72.

Schwalm, C. R., S. Glendon, and P. B. Duffy. 2020. RCP8.5 tracks cumulative CO2 emissions. Proceedings of the National Academy of Sciences of the United States of America 117:19656–19657.

Shelton, A. O., and M. Mangel. 2011. Fluctuations of fish populations and the magnifying effects of fishing. Proceedings of the National Academy of Sciences of the United States of America 108:7075–7080.

Silvy, Y., E. Guilyardi, J. B. Sallée, and P. J. Durack. 2020. Human-induced changes to the global ocean water masses and their time of emergence. Nature Climate Change 10:1030–1036.

Stark, J. W. 2007. Geographic and seasonal variations in maturation and growth of female Pacific cod (Gadus macrocephalus) in the Gulf of Alaska and Bering Sea. FISHERY BULLETIN 105:396–407.

Stige, L. C., M. E. Hunsicker, K. M. Bailey, N. A. Yaragina, and G. L. Hunt. 2013. Predicting fish recruitment from juvenile abundance and environmental indices. Marine Ecology Progress Series 480:245–261.

Vehtari, A., A. Gelman, and J. Gabry. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing 27:1413–1432.

Walsh, J. E., U. S. Bhatt, J. S. Littell, M. Leonawicz, M. Lindgren, T. A. Kurkowski, P. A. Bieniek, R. Thoman, S. Gray, and T. S. Rupp. 2018a. Downscaling of climate model output for Alaskan stakeholders. Environmental Modelling and Software 110:38–51.

Walsh, J. E., R. L. Thoman, U. S. Bhatt, P. A. Bieniek, B. Brettschneider, M. Brubaker, S. Danielson, R. Lader, F. Fetterer, K. Holderied, K. Iken, A. Mahoney, M. McCammon, and J. Partain. 2018b. The high latitude heat wave of 2016 and its impacts on Alaska. Bulletin of the American Meteorological Society 99:S39--S43.

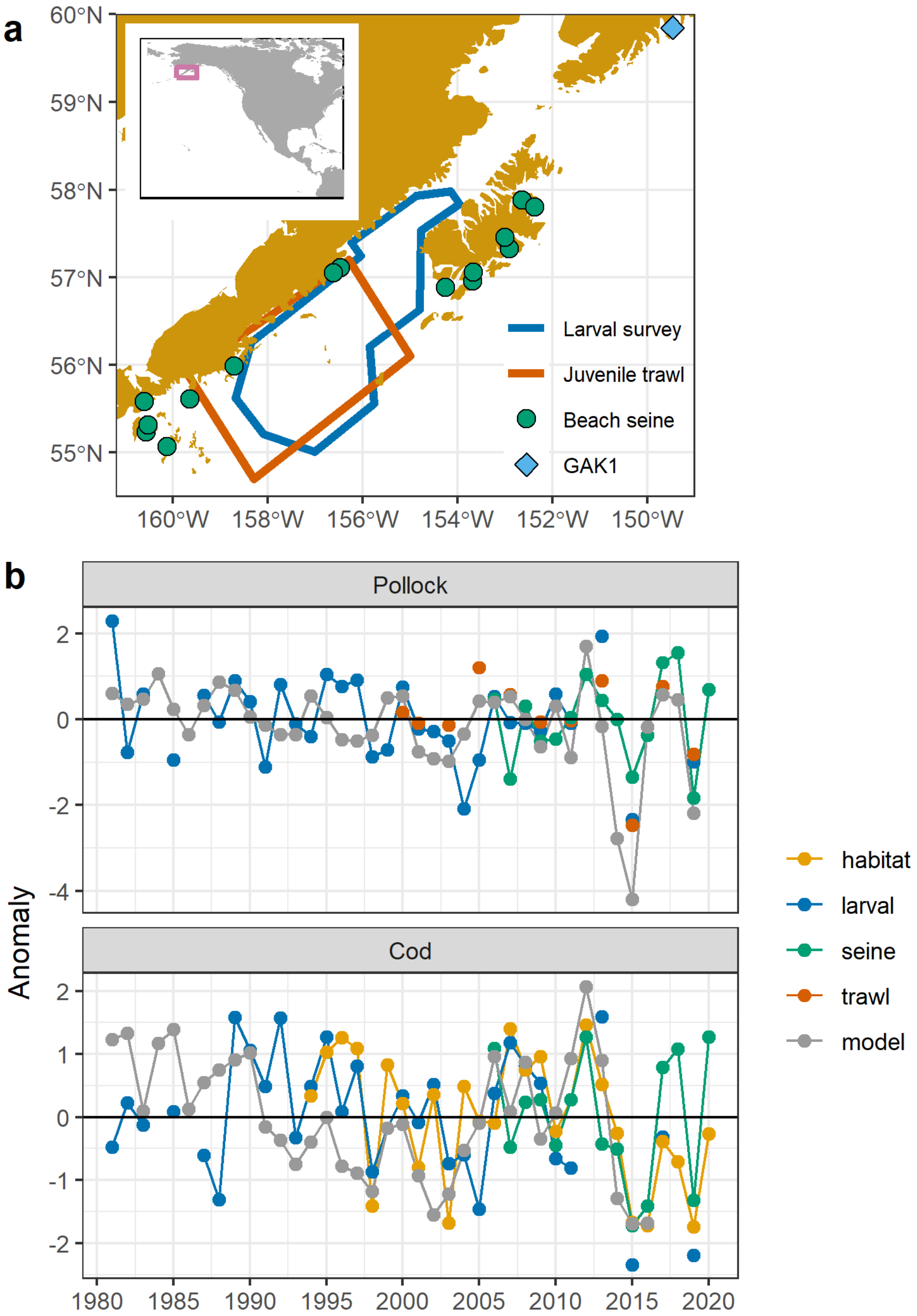
Williams, J. W., and S. T. Jackson. 2007. Novel climates, no-analog communities, and ecological surprises. Frontiers in Ecology and the Environment 5:475–482.

Wilson, M. T., and N. Laman. 2021. Interannual variation in the coastal distribution of a juvenile gadid in the northeast Pacific Ocean: The relevance of wind and effect on recruitment. Fisheries Oceanography 30:3–22.

Wolkovich, E. M., B. I. Cook, K. K. McLauchlan, and T. J. Davies. 2014. Temporal ecology in the Anthropocene. Ecology Letters 17:1365–1379.

Wood, S. N. 2003. Thin plate regression splines. Journal of the Royal Statistical Society. Series B: Statistical Methodology 65:95–114.

Zuur, A. F., I. D. Tuck, and N. Bailey. 2003. Dynamic factor analysis to estimate common trends in fisheries time series. Canadian Journal of Fisheries and Aquatic Sciences 60:542–552.



581

582 Fig. 1. Study system. a) Study site. b) Age-0 field data and stock assessment model recruitment estimates

583 for each species.

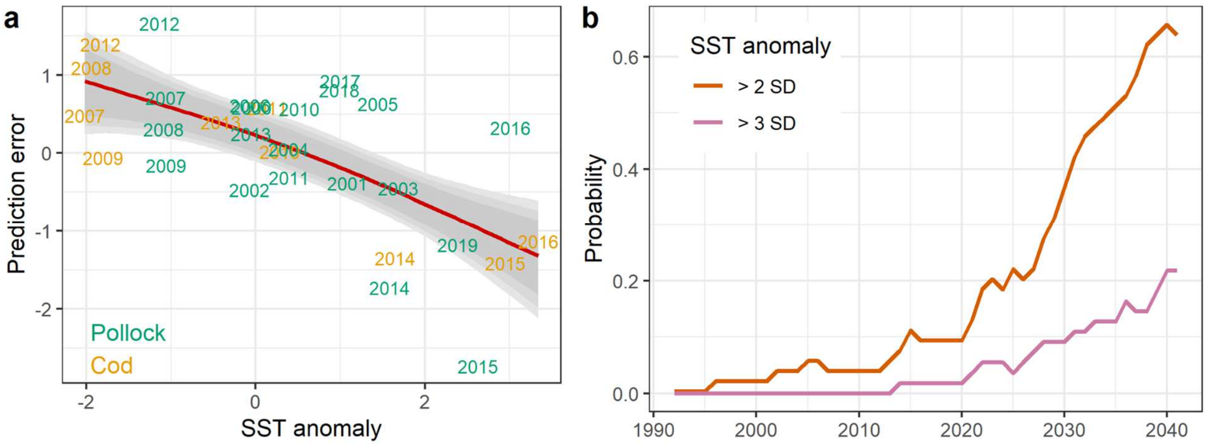
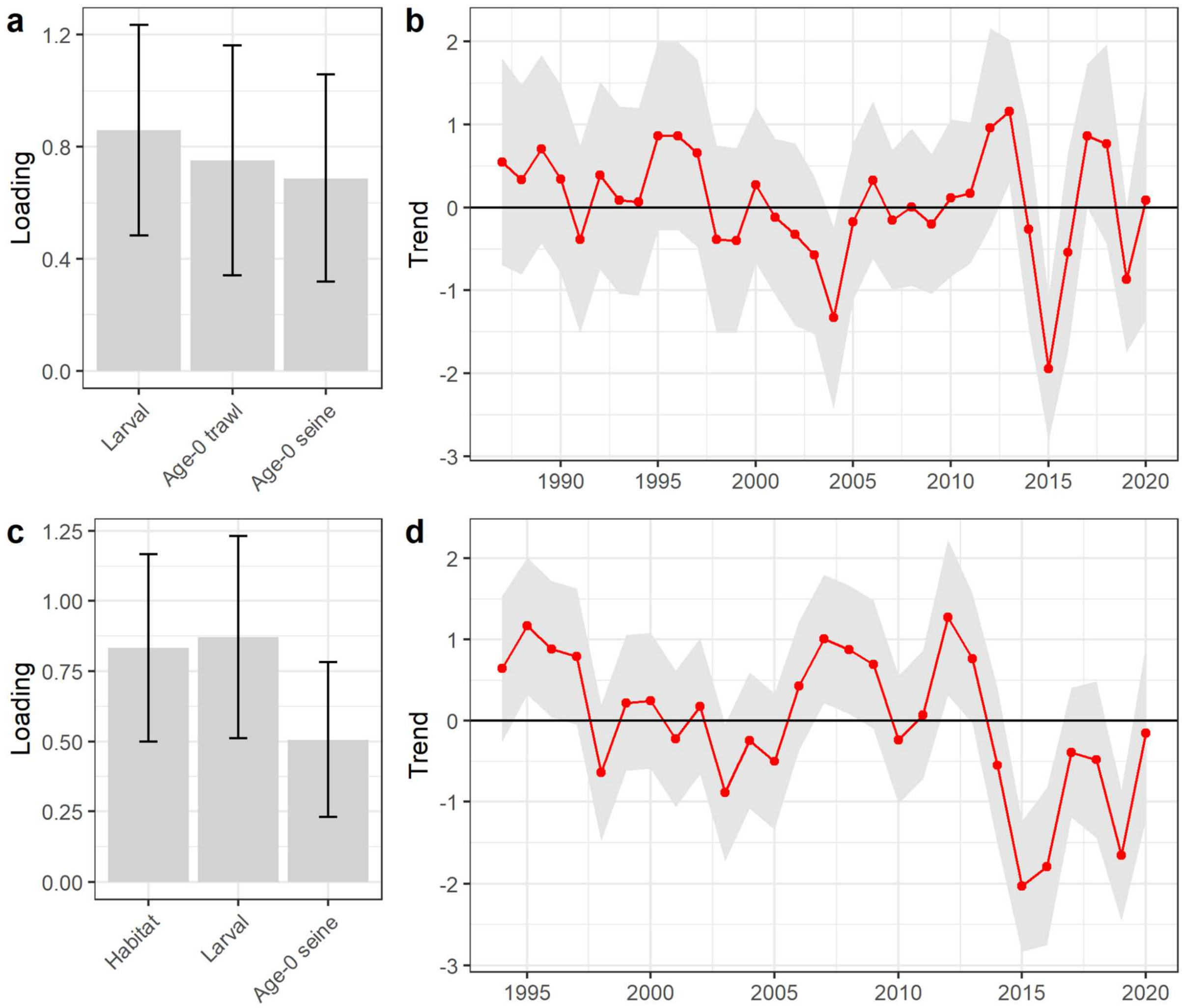


Fig. 2. Climate change and prediction error for recruitment at age-0. a) Prediction error for both species combined as a function of annual SST anomaly (posterior means with 80 / 90 / 95% CI). b) Probability of > 2 SD and > 3 SD SST anomalies (relative to 1987-2016 base period) from downscaled CMIP5 models under emissions scenario RCP8.5: mean annual probabilities for five models, smoothed with 11-year running mean (redrawn from data in Walsh et al. 2018).



590

591 Fig. 3. Loadings and shared trend from DFA models for shared trend in recruitment. a-b) Time series

592 loadings and shared trend for (a-b) pollock, (c-d) cod.

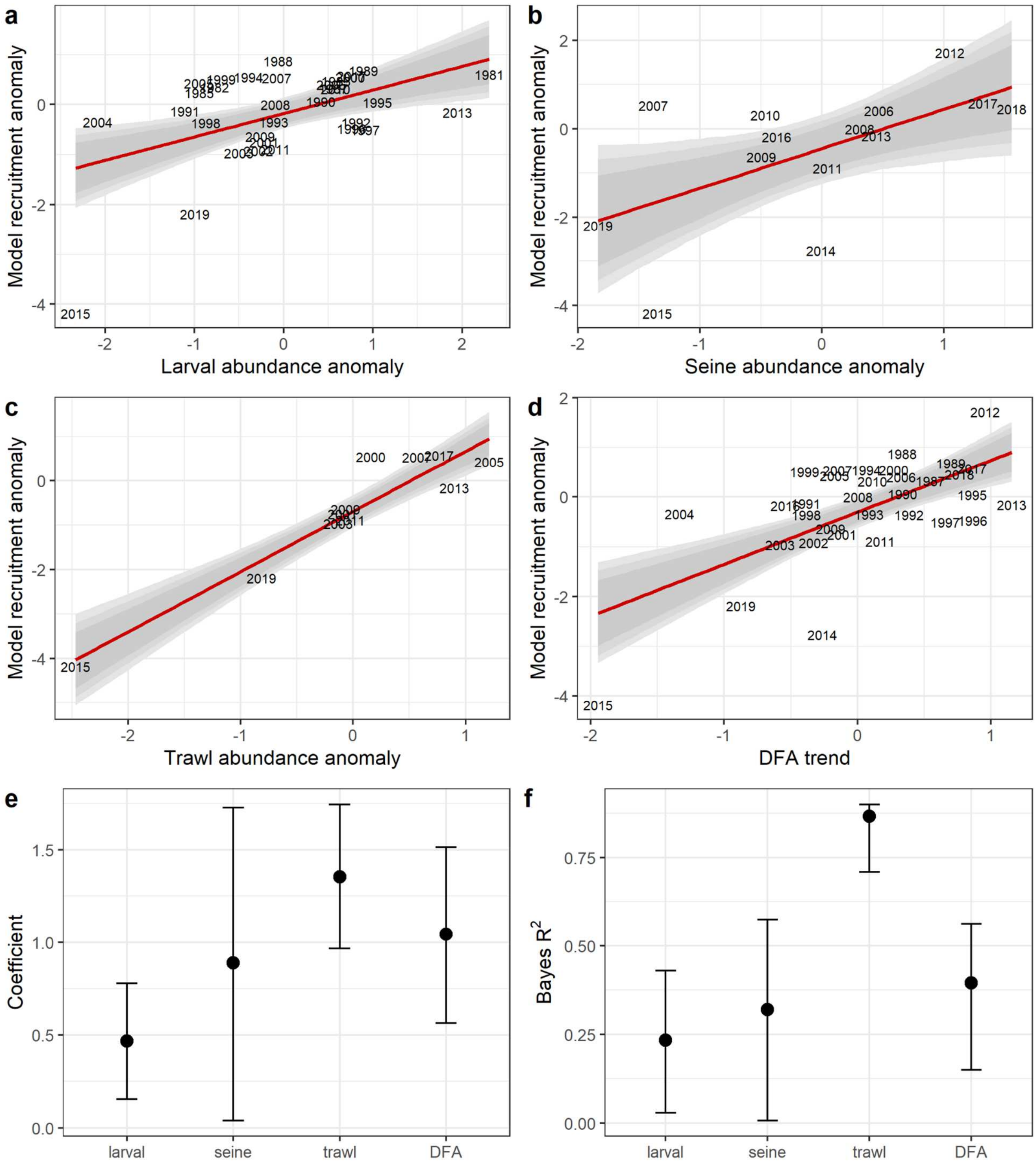


Fig. 4. Predicting pollock model recruitment estimates from age-0 field observations. Bayesian regression (80 / 90 / 95% CI) of model estimates on a) larval abundance, b) juvenile seine abundance, c) juvenile trawl abundance, d) DFA for all three time series. e) Estimated regression coefficients (95% CI) for each model, f) Bayes R^2 (95% CI) for each model.

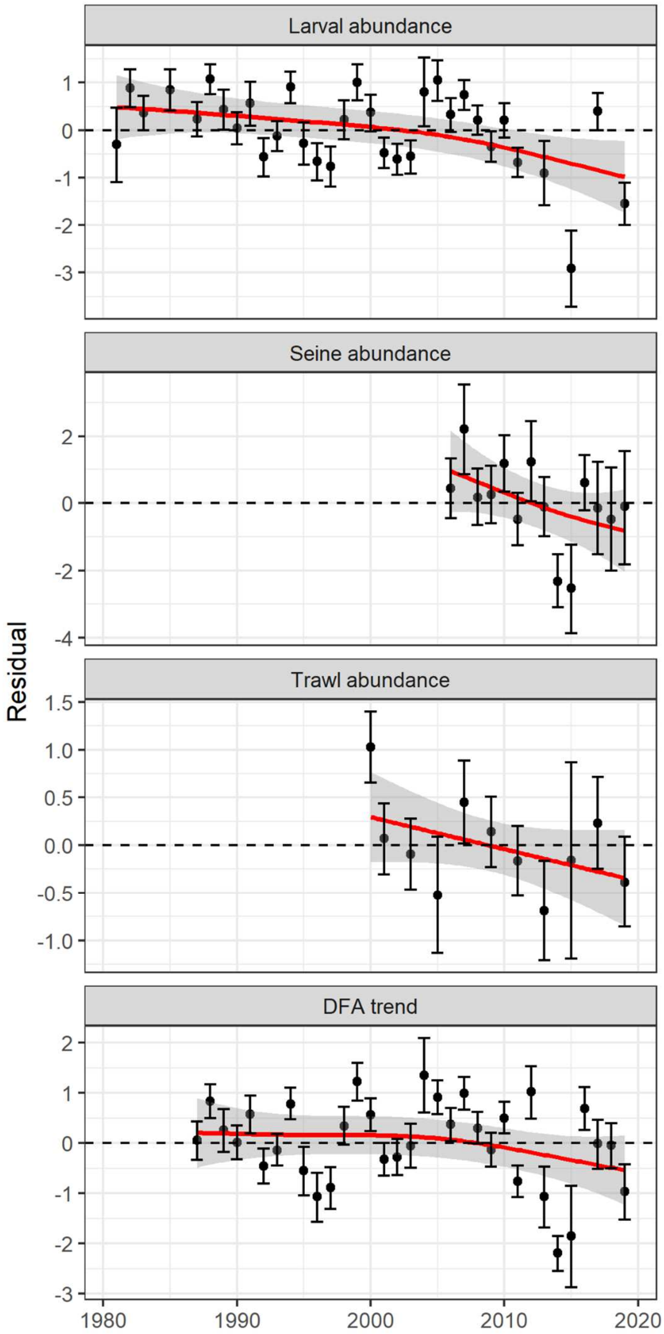


Fig. 5. Residual time series for Bayesian predictions of pollock stock assessment model recruitment estimates: a) larval abundance, b) seine abundance, c) trawl abundance, d) DFA trend. Residuals are posterior medians and 95% CIs, and fitted regressions are from Generalized Additive Models, with 95% CIs, calculated under the assumption that years are independent observations.

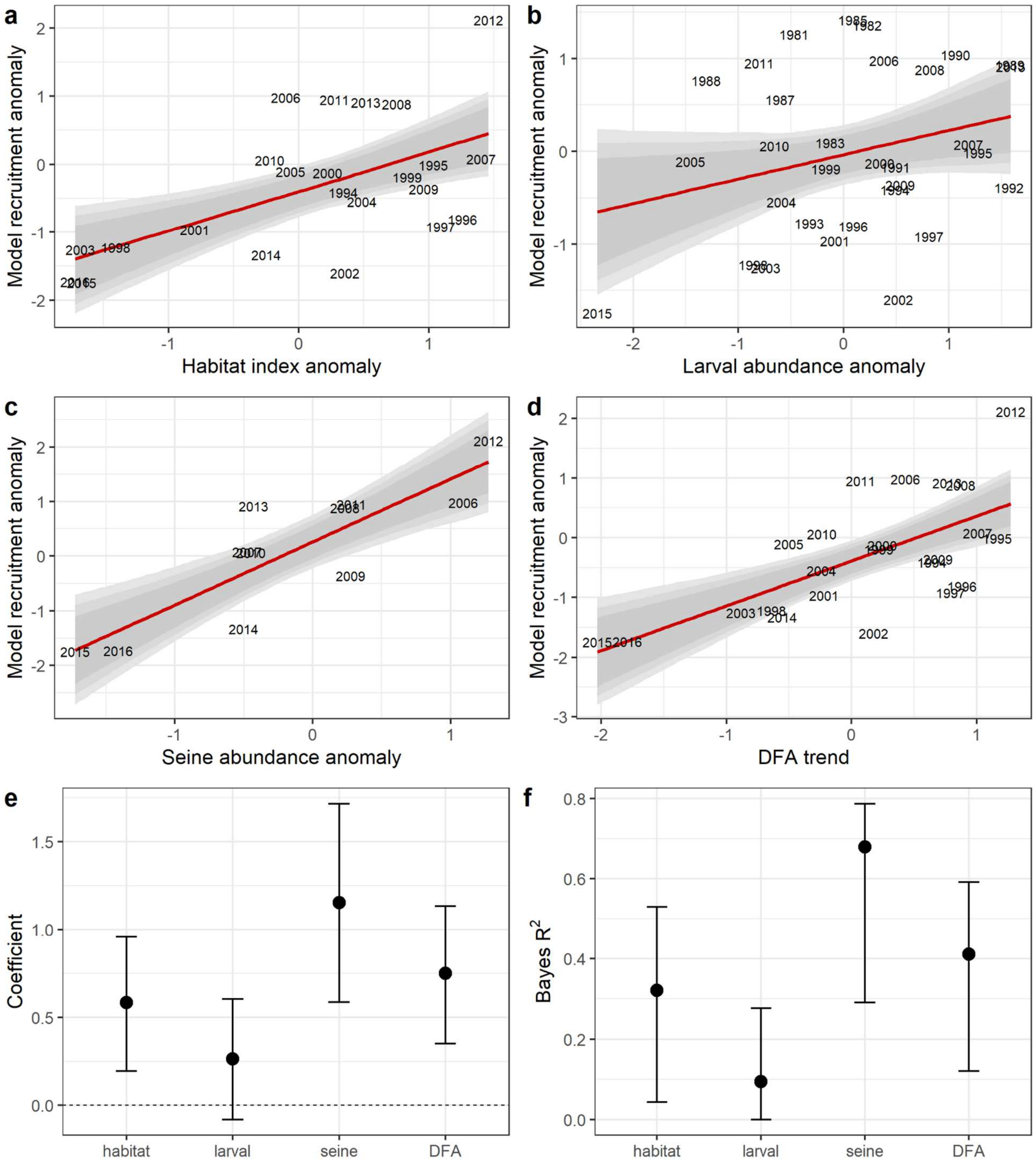


Fig. 6. Predicting cod model recruitment estimates from age-0 field observations. Bayesian regression (80 / 90 / 95% CI) of model estimates on a) spawning habitat index, b) larval abundance, c) juvenile seine abundance, d) DFA for all three time series. e) Estimated regression coefficients (95% CI) for each model, f) Bayes R^2 (95% CI) for each model.

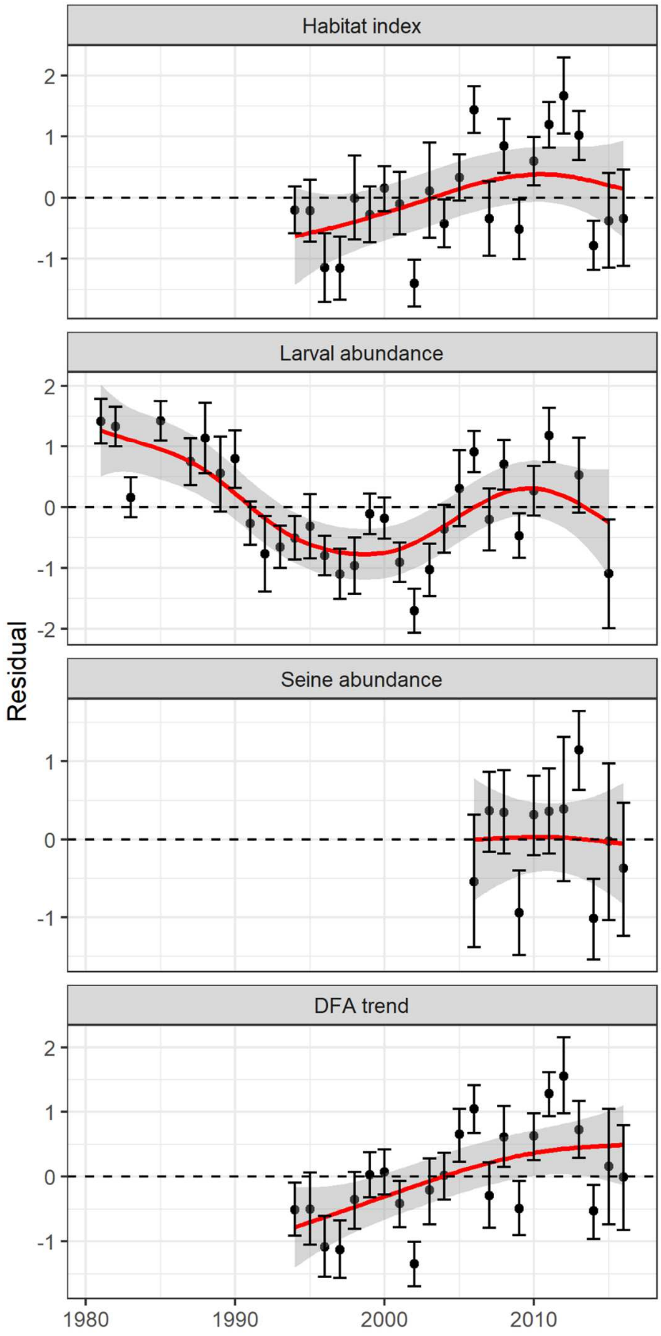


Fig. 7. Residual time series for Bayesian predictions of cod stock assessment model recruitment estimates: a) spawning habitat index, b) larval abundance, c) seine abundance, d) DFA trend. Residuals are posterior medians and 95% CIs, and fitted regressions are from Generalized Additive Models, with 95% CIs, calculated under the assumption that years are independent observations.