Sea-level index of recruitment variability improves assessment model performance for sablefish *Anoplopoma fimbria*

Tolimieri N[1] and Haltuch MA[2]

[1]Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Boulevard E, Seattle, Washington 98112, USA

[2]Fishery Resource Analysis and Monitoring Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Boulevard E, Seattle, Washington 98112, USA

Corresponding Author: nick.tolimieri@noaa.gov

Short Title: Sea-level and sablefish stock assessment

Abstract

Environmental recruitment indices may improve the precision of stock assessments, allow hindcasting, and aid in near-term forecasting. We used Bayesian dynamic factor analysis (DFA) to find common trends in sea level from 16 tide gauges spanning the US West Coast. We then used these DFs as predictors of sablefish *Anoplopoma fimbria* recruitment deviations from the 2021 assessment. We evaluated the ability of the resulting northern sea-level index (north of Cape Mendocino, ~40° N) to inform recruitment estimates and its impacts on assessment model predictions by running two hindcast stock assessment models: 1) a catch-only model, which assumed average recruitment from the stock-recruit relationship, and 2) a catch plus sea-level model. In both cases, survey data were removed from 2011 forward. The model including sea-level index captured the observed increase in stock biomass from 2016 onwards, while the catch only model did not, predicting a continued biomass decline. This work provides evidence of the potential to improve forward-looking stock projections by better capturing stock trends, providing an advance over average recruitment assumptions.

Key words: Bayesian dynamic factor analysis, sea-surface height, environmental variability,

Introduction

Fisheries managers face the combined problem that recruitment is seldom average due to variation in environmental drivers, and that cohort strength is often poorly estimated until the cohort is several years old and well sampled by the fishery or fishery independent surveys. As a result, multiple years of data are often required to produce precise estimates of recruitment. Thus, it is difficult to provide accurate, short-term forecasts of cohort strength and stock biomass, and scientists and managers must often wait several years for sufficient data to support good estimates of cohort strength. Likewise, hindcasting to periods of low data availability or poor data quality often must rely on an assumption of average recruitment from the spawner-recruit relationship, which is unlikely to be accurate—the environment influences recruitment and productivity of populations of many marine fishes. For species with weak stock-recruitment relationships, the inclusion of environmental recruitment indices in stock assessments may provide a route towards improving model precision, allowing hindcasting during periods of low data availability, and aiding in near-term forecasting (Stige et al. 2013; du Pontavice et al. 2022).

Sablefish *Anoplopoma fimbria* inhabit waters along the west coast of North America from Baja California through Alaska and extend west (and south) to Japan (Hart 1973; Allen and Smith 1988; Johnson et al. 2016). While managed as three separate stocks in the eastern Pacific (Alaskan, British Columbian, and US West Coast), sablefish genetic analyses have not found strong population structure, which suggests a single panmictic genetic population in the northeastern Pacific from California to Alaska (Jasonowicz et al. 2017). Additionally, regional trends in recruitment and spawning stock biomass estimates from - stock assessment models (Alaska, British Columbia, and the West Coast) demonstrate some synchrony across the three management regions (Fig. S1), although this synchrony has broken down in recent years (Fenske et al. 2019). However, the sablefish do show evidence of spatial structure in growth (Kapur et al. 2020) and maturity (Head et al. 2014), likely due to environmental differences across its range.

Sablefish is one of the most valuable stocks in the region. For example, in 2018, West Coast fisheries landed 5275 metric tons of sablefish with an ex-vessel value of $24.7 million USD (Haltuch et al. 2019b). However, the West Coast stock was estimated to have been in decline since the mid 1970's, due to fishing pressure compounded by a period of lower than expected recruitments, only recently experiencing an increasing trend due to a few large recruitment events (Johnson et al. 2016; Haltuch et al. 2019b; Kapur et al. 2021). To better understand the persistent stock decline and recent increase, there has been a substantial focus on examining environmental predictors of recruitment for the West Coast stock, with the goal of improving the weak stock-recruitment analytical relationship (Fig. 1) (Schirripa and Colbert 2006; Tolimieri et al. 2018).

Tolimieri et al. (2018) used output from the Regional Ocean Modeling System (ROMS) physical oceanographic model for the California Current Ecosystem (Neveu et al. 2016) to test life-history based, mechanistic hypotheses for potential environmental recruitment drivers (1980-2010, 40-48° N). Residuals from the stock-recruitment curve (indicating deviations from modeled median recruitment) were positively correlated with colder conditions during the spawner preconditioning period, warmer water temperatures during the egg stage, stronger cross-shelf transport to near-shore nursery habitats during the egg stage, stronger long-shore transport to the north during the yolk-sac stage, and cold surface water temperatures during the larval stage. While informative and often times more available than some observational data streams (du Pontavice et al. 2022), using ROMS predictors has several drawbacks including the need to update the ROMS annually, limited historical time frames for available outputs (e.g., 1980-2010), and potential discontinuities as models are updated and data inputs change (Tolimieri et al. 2018; Haltuch et al. 2019a). These challenges make it difficult to conduct analyses, such as hindcasting, to better understand historical biomass prior to the commencement of heavy exploitation, or now- or near-term forecasting for catch-only stock assessment projections or management strategy evaluations.

In addition to the ROMS-recruitment research, there is an established relationship between sea level and sablefish recruitment; recruitment is negatively correlated with sea level north of Cape Mendocino (Schirripa and Colbert 2005; Schirripa and Colbert 2006; Schirripa 2007; Stewart et al. 2011; Johnson et al. 2016), a known biogeographic barrier (Tolimieri 2006; Tolimieri and Levin 2006). Changes in sea level serve as a proxy for large-scale climate forcing that drives regional changes in alongshore and cross-shelf ocean transport (Reid and Mantyla 1976; Chelton and Davis 1982). Lower sea level in the north correlates with stronger upwelling and alongshore surface flow to the south (Reid and Mantyla 1976; Chelton et al. 1982; Chelton and Davis 1982; Chelton 1984). Low sea level in the northern California Current Ecosystem is also related to a stronger alongshore sea-level gradient (higher in the south, lower in the north). This alongshore sea-level/pressure gradient drives a stronger poleward deep current that tends to be strongest between about 100 and 500m, although poleward flows extend deeper (Connolly et al. 2014). Southerly transport of surface waters brings fatty acid-rich northern copepods into the California Current (Chelton et al. 1982; Keister et al. 2011), which are an important food resource for sablefish and many other consumers

(Grover and Olla 1987; Mcfarlane and Beamish 1992; King et al. 2000). Low sea level is also associated with northern source waters that are more “minty”, cooler water with higher dissolved oxygen (Schroeder et al. 2019), resulting in higher productivity from upwelling. Mechanistic modeling using ROMS predictors (Tolimieri et al. 2018) suggests that northerly transport at depths around 1000 m (which mirrors deep transport under low sea-level conditions in the north) leads to stronger year-class strength by bringing yolk-sac larvae to the north where they can encounter these northern copepods once the sablefish larvae rise to surface waters and begin feeding. Variability in sea level has also been linked to the abundance of pelagic young-of-the-year stages of rockfish (*Sebastes* spp.) in the California Current, where low sea level is associated with equatorward flow and the predominance of cooler, oxygen-rich Pacific Subarctic Upper Water (Ralston et al. 2013; Schroeder et al. 2019) compared to more southern waters (Schroeder et al. 2019).

While the relationship between sablefish recruitment and sea level has been weaker than the relationship with the five ROMS variables, the sea-level data are valuable because they cover a longer, continuous time span than ROMS outputs, are updated reliably in quasi-real time for multiple sites along the US West Coast, and the relationship has withstood repeated testing during the stock assessment process (Schirripa and Colbert 2005; Schirripa and Colbert 2006; Schirripa 2007; Stewart et al. 2011; Johnson et al. 2016). Thus, the temporal availability of these data make them viable as an index of recruitment for both fore- and hindcasting.

Stock assessments for the West Coast groundfish fishery use the Stock Synthesis model (Methot and Wetzel 2013) to integrate data from multiple sources including fishery independent data such as abundance indices, size, and age data derived from the West

Coast Groundfish Bottom Trawl Survey (Bradburn et al. 2011), and fisheries catch and bycatch data from commercial and recreational fisheries. Previous sea-level analyses have selected individual tide-gauge locations (Schirripa and Colbert 2006) based on the strength of the resulting relationship with recruitment or averaged variation from measurements at several sites on the northern West Coast of the US (Schirripa et al. 2009; Stewart et al. 2011). This sea-level index is not spatially integrative, and therefore may not be representative of the full coast. We first used a sea-level index derived from dynamic factor analysis (DFA) in the 2019 sablefish assessment (Haltuch et al. 2019b), which, as a benchmark assessment, went through rigorous review and acceptance of both the data and model to be used for fishery management. The 2021 update of the 2019 assessment (Kapur et al. 2021) permits limited model changes with updated data and is currently the basis for managing U.S. West Coast sablefish fisheries.

With the aim of improving model precision, allowing hindcasting during periods of low data availability, and aiding in near-term forecasting, we (1) use Bayesian dynamic factor analysis (DFA, Ward et al. 2021) to look for common trends in the sea-level time series from 16 locations covering the full extent of the US West Coast from San Diego, CA north to Neah Bay, WA. Next, (2) we use the resulting dynamic factors to predict recruitment deviations for 1975-2020, derived from the 2021 sablefish stock assessment (Kapur et al. 2021). Finally, (3) we use a retrospective or hindcast analysis based on the 2021 sablefish assessment to assess whether the sea-level index provides enough information to improve prediction of modeled recruitment estimates within the stock-assessment model when only commercial catch data are available to the assessment.

Materials and methods

Life-history

Sablefish are bathy-demersal, inhabiting deep waters (175 – 2740 m) along the west coast of North America from Baja California through Alaska and extending west and south to Japan (Hart 1973; Allen and Smith 1988; Johnson et al. 2016). Along the US West Coast, spawning occurs from December to March (peak February) at the edge of the continental shelf at depths greater than 300 m (Mason et al. 1983; Boehlert and Yoklavich 1985; Kendall and Matarese 1987; Hunter et al. 1989; Moser et al. 1994). Eggs are buoyant and rise to 200-300 m in the water column (but can be found as deep as 480 m). After approximately 12-17 days, the eggs hatch (Mason et al. 1983; Kendall and Matarese 1987; Mcfarlane and Beamish 1992; Moser et al. 1994), and yolk-sac larvae sink to 1000-1200 m where they are found between February and May. Larvae move to surface waters by 40 days post-hatch and are encountered from the 500-m isobath out to 150 nautical miles (277 km) during the same February to May, as spawning is prolonged (Brock 1940; Mcfarlane and Beamish 1992; Moser et al. 1994). Pelagic juveniles also stay in these surface waters and are present from April through November (Mitchell and Hunter 1970; Kendall and Matarese 1987). Age-0 recruits settle to the benthos between August and November with most fish settling to habitats 250 m or shallower.

Sea-level data

We used time series of monthly mean sea level from 16 tide gauges spanning the US West Coast (NOAA Tides and Currents https://tidesandcurrents.noaa.gov/sltrends/, Fig. 2). Specifically, we used the data for inter-annual variation, which have the average seasonal cycle and linear trend removed, allowing us to index the inter-annual variation in environmental and oceanographic drivers that may drive inter-annual variation in recruitment. We then calculated the mean spring sea level (April to June, Fig. S2), when multiple life-stages are in the water column (Tolimieri et al. 2018). This period is consistent with the timing of previous analyses of sea level height and sablefish recruitment (Schirripa 2007; Schirripa et al. 2009; Stewart et al. 2011).

Dynamic factor analysis (DFA)

We used Bayesian dynamic factor analysis (Ward et al. 2021) to identify common trends in sea level among the 16 locations and to develop potential environmental indices of sablefish recruitment. DFA is a time-series analog for principal components analysis that estimates common trends in multiple time series while accounting for autocorrelation and allowing different observational error structures (Holmes et al. 2021). Importantly, DFA can handle missing data and time series of different lengths (Zuur et al. 2003b; Zuur et al. 2003a). We included the mean spring sea level for the 16 tide gauge stations for the years 1925-2020 in the DFA analysis.

Haltuch et al. (2019b) used a non-Bayesian DFA framework to evaluate model structure and evaluated models allow 1-5 dynamic factors and different error structures (diagonal and equal, diagonal and unequal). Based on that analysis, we ran a single Bayesian DFA ('bayesdfa' package in R, R Core Team 2021; Ward et al. 2021) to estimate 95% credible intervals for the resulting dynamic factors in order to provide uncertainty estimates for inclusion in the stock assessment portion of the analysis. Including this uncertainty is important for use as an index in stock analysis because it allows one to evaluate how uncertainty in the index impacts output from the assessment model. We used the same model parameters as the best-fit model from the non-Bayesian with five dynamic factors, and a diagonal and unequal variance covariance matrix. We used three chains and 3000 iterations following a 1500 burn-in period. We standardized the sea-level data by subtracting the mean and dividing by the standard deviation prior to analysis, which is a standard approach for DFA (Holmes et al. 2021). Note the order of the DFs is not indicative of explained variance as in principal components analysis.

Modeled recruitment deviations

Estimates of the $\log_e$ recruitment deviations from the 2021 sablefish stock assessment (Kapur et al. 2021) were used in the following analyses. Log_e recruitment deviations estimated from the stock assessment provide model-based, annual estimates of the difference between each year's recruitment and the fitted stock-recruit relationship that provides estimates of the median, deterministic recruitment expected in a given year. The sablefish stock assessment assumes a Beverton-Holt stock-recruitment function with $\log_e$ recruitment deviations that vary annually, due to processes not modeled in the stock assessment, and undergo bias correction (Methot and Taylor 2011; Kapur et al. 2021).

Sea level – recruitment model fitting

To determine whether sea level functioned as a predictor of sablefish recruitment, we used the $\log_e$ bias-corrected recruitment deviations around the Beverton-Holt stock-recruitment curve from the 2021 sablefish stock assessment (Kapur et al. 2021) as the response variable in general linear models using five DFs as predictor variables (hereafter "recruitment model(s)"). We limited the time period to 1975 - 2020 because of a paucity of size and age data prior to 1975 and because assessment-based recruitment deviations and sea-level data were both available through 2020 (Kapur et al. 2021). This time period is broader than the 1980-2010 analysis of the ROMS variables and sablefish recruitment (Tolimieri et al. 2018), and once developed, the index could, in theory, be used to hindcast farther back in time than 1975 to inform recruitment in earlier time periods. We included both linear and quadratic terms in the model fitting but required that any model including a quadratic term (e.g., $DF1^2$) also include its linear counterpart (DF1). We then ran all possible combinations of the five DFS and used ΔAICc to compare candidate models (Burnham and Anderson 1998). We examined all candidate models (ΔAICc < 2.0) and identified the one with the fewest parameters as the best-fit recruitment model. While it would be worthwhile to occasionally re-evaluate the relationship, the expectation would be to calculate and use the resulting index in the assessment – not re-run all of the model selection each assessment.

We ran an array of additional tests to validate the recruitment model results and fit (see Supplementary Material) following Tolimieri et al. (2018) and Haltuch et al. (2019b). Here, we highlight several of these tests. First, we refit the best-fit model to the recruitment deviations for 1975-2015 (but using the sea-level index derived from the 1975-2020 DFA) and then used that model to predict recruitment for 2016-2020 to determine how consistently the model forecast performed relative to the full 1975-2020 best-fit model. Note because the refit model excludes the recruitment data for 2016-2020, the coefficients, and therefore predictions, may differ between the best-fit model using all the data and the subsetted model. Second, we refit the best-fit model to 1975-2015 and then predicted recruitment deviation for the next year 2016. We then iteratively added a year to the refitting and predicted the next year's recruitment deviation. These two approaches address the ability of the sea-level index to inform future recruitment over different periods (5 years or one year at a time) based on the relationship estimated over an earlier period. We also conducted a jackknife analysis dropping one year at a time and refitting the model to determine whether individual years had strong effects on the model predictions and to estimate bias. Finally, to determine whether the terms included in the best-fit model might differ over a shorter time period, we reran the entire model selection process using recruitment data for 1975-2015 only (but using the DFA results for 1975-2020 but including only 1975-2015). See Supplementary Material for additional model validation.

Stock assessment hindcast

We used the 2021 sablefish assessment (Kapur et al. 2021) to conduct the hindcast analyses. The 2021 stock assessment used the sea-level index as an index of recruitment deviations in the same manner in which a survey index of abundance would be used in a stock assessment model (Methot and Wetzel 2013, Methot et al. 2022). The 95% credible intervals from the DFA analysis were used in the stock assessment model to characterize the annual variability in the sea-level index. The relationship of the sea-level index with the recruitment deviations was assumed to be proportional and was estimated by a single time invariant parameter (Methot and Wetzel 2013, Methot et al. 2022). The stock assessment model also estimated an additional standard error parameter that was an additive constant added to the input standard deviation of the survey variability (Methot and Wetzel 2013, Methot et al. 2022). First, we evaluated the impact of the sea-level index on the model results (time series of spawning biomass, recruitment deviations, fraction of the unfished spawning biomass) by comparing the results of the 2021 assessment model (base model plus sea level) to the same model without the sea-level index (base model). Differences between the two sets of model results were minimal (see Results), indicating that recruitment deviations were largely informed by survey age data (an expected outcome) and provide context for the use of this model for hindcast comparisons.

Next, we ran two hindcast models to determine whether sea level could predict deviations in recruitment without fishery dependent and independent data informing the population dynamics. Both hindcast models removed all fishery dependent and survey data from 2011 forward, except for commercial catch data, and fixed all selectivity parameters. The catch-only hindcast model also removes the sea-level index, while the catch plus sea-level hindcast model retains the sea-level index. These two hindcast models treat the years 2011 to 2020 as a projection period, and span a similar period as model forecasts provided for management. We then compare these two hindcast models to the full stock assessment, which represents the 'true' state.

We evaluated the value of including sea-level recruitment index in three ways. First, we evaluated the ability of the each hindcast (catch plus sea-level versus the catch-only) to capture trends in stock size observed in the full 2021 assessment, specifically for the 2011-2020 period. Second, we compared the number of years that each hindcast model captured the direction of change in the recruitment deviations. Third, we calculated the percent absolute difference (ARD) (see Haltuch and Punt 2011, equation 17) for each hindcast model (catch-only and catch plus sea level) and year from 2011 to 2019, resulting in a time series of nine different percent-ARDs for each time series of recruitment deviations, recruitment, spawning biomass, and fraction of the unfished spawning biomass. Then, for each hindcast model, the median of the nine annual percent ARDs for recruitment deviations and the mean of the nine annual percent ARDs for each time series of recruitment, spawning biomass, and fraction of unfished spawning biomass are reported as single summary statistics where higher values indicate poorer performance and values close to zero indicate better performance. All years included in the median and mean calculations are equally weighted. Note, the 2020 estimates are excluded here because there are no survey data for 2020, and the sea-level index is the primary source of data informing estimation of recruitment deviations for both models in 2020.

Results

Dynamic factor analysis: sea-level trends

The five dynamic factors (Fig. 3) had a generally good fit to the data (Fig. S3), and factor loadings (Fig. 4) identified three broad latitudinal trends. DF1 (hereafter, northern sea-level index) characterized variation in sea level from North Spit (approximately Cape Mendocino) to the north (positive loadings, Fig. 4). DF3 indexed variation in sea level among mid-latitude locations from approximately Crescent City to Monterey or Port San Luis, while DF4 included more southerly locations from Santa Monica to San Diego. The other two DFs did not show strong spatial trends.

Model selection: predicting recruitment deviations

Model selection evaluating the number of sea-level DFs to include for predicating recruitment deviations identified two recruitment models had ΔAICc values less than 2.0. Both recruitment models included the northern sea-level index (DF1) indicating that oceanographic processes in the northern portion of the West Coast were important for determining recruitment. Model 1 included only the northern sea-level index, while Model 2 also included the southern sea-level index, DF4. Model 1 had the lowest AICc and fewest parameters, so we selected it as the best-fit model.

The best-fit recruitment model (Model 1: recruitment deviations ~ DF1) explained 15% of the variation in the recruitment deviations from 1975-2020 (Fig. 5). Recruitment deviations were negatively correlated with the northern sea-level index (Table 1, Fig. 6) and, therefore, negatively correlated with sea level north of approximately Cape Mendocino. The low predictive power ($r^2 = 0.15$) appears to be due to the model failing to predict lower than expected recruitments (Fig. 5), especially in 2006, 2007, and 2009, and to changes in recruitment estimates between the 2019 benchmark stock assessment (Haltuch et al. 2019b), and the 2021 update stock assessment (Kapur et al. 2021). It is common for recruitment estimates to vary between models, particularly during periods with recruitment estimation poorly, or not, informed by data. Such periods often include early model periods with little to no age-composition data, and the last few years of assessments where there are few data on recruitments entering the population from surveys due to size-based catchability (Bradburn et al. 2011; Tolimieri et al. 2020). The sablefish assessments can estimate large changes in recruitment estimates during the 1960s and 1970s due to a lack of informative age data for this period, resulting in smaller shifts to subsequent recruitment estimates. In other cases, the model under-predicted strong recruitments or over-predicted weak recruitments even though it did predict peaks or lows in those years. However, the data quality of the recruitment time series generally increases with time as more information enters the stock assessment model and recruitment deviations are better estimated. The amount of variation in recruitment explained by the northern sea-level index (DF1) was low when considering just the early portion of the time series from 1975-2002 ($r^2 = 0.07$, $p = 0.15$) where survey data were limited. However, from 2003 to 2020, when the assessment was informed by an annual fishery-independent trawl survey (Keller et al. 2017), the fit was much better ($r^2 = 0.28$, $p = 0.02$).

Model testing and validation showed the best-fit recruitment model to be consistent and stable (Fig. 5, see also Supplementary Material, Table S1, Fig. S4-S6). Refitting the recruitment deviations for 1975-2015 and then predicting 2016-2020 differed little from the 1975-2020 model results. Likewise, fitting 1975-2015 and then stepping forward one year at a time was also consistent with the 1975-2020 model. Finally, removing individual years and refitting the best-fit model (jackknife resampling) had little effect on the model fit (median $r^2 = 0.15$, 95% C.I. = 0.12– 0.19, Fig. S4). Recruitment deviations were consistent with the best-fit model, with only a minor difference when excluding 1993. Limiting the analysis to the 1975-2015 period and re-running the entire model selection process produced the same best-fit model, which included only DF1, which indexed sea level north of Cape Mendocino.

Stock assessment hindcast

Removing the sea-level index from the 2021 assessment while retaining all other data had only a minor impact on the model outputs (see Supplementary Material). The recruitment estimates from approximately 1950 to approximately 1975 were smoother, and a major recruitment peak shifted earlier in the time series, resulting in slightly earlier increases in sablefish biomass in the late 1960s than when the sea-level index was included (Fig. S7). However, removing the sea-level index from the 2021 stock assessment did not have strong effects on the assessment results from 1975 onwards when the assessment is increasingly well informed by age data, and because the age data and sea-level index provide similar information on recruitment. This result supports using the 2021 stock assessment model as the basis for the hindcast model runs.

In the 2021 stock assessment, sablefish spawning biomass increased from 2016 to 2021 after a long period of decline (Fig. S7). The catch plus sea-level hindcast for 2011 onward was able to capture this increasing trend in stock size, but the catch-only hindcast showed persistent stock decline due to the inability to capture above-average recruitments in 2013, 2015, and 2016 (Tables S2 & S3, Fig. 7). Over the years 2011 to 2019, the catch plus sea-level hindcast captured the direction of change in the recruitment deviations, in comparison to the best estimates from the 2021 stock assessment, in six out of nine years (2012, 2013, 2015, 2016, 2018, 2019) (Tables S2 & S3, Fig. 7). Four years underestimated the magnitude of change (2012, 2015, 2016, 2018), two years were small overestimates of positive deviations (2013, 2019). The catch plus sea-level hindcast was also able to capture recruitment deviations away from the long-term average recruitment deviations, although the larger recruitment deviations were generally underestimates compared to the 2021 stock assessment. In 2017, the catch plus sea-level hindcast did not capture the direction of change in recruitment deviations, underestimating a recruitment deviation above the long-term average. Percent median absolute relative differences for recruitment deviations from the catch-only and catch plus sea-level hindcasts were 103%, and 43%, respectively, with the lower value indicating greater agreement with the 2021 stock assessment. Thus, the catch plus sea-level hindcasts were better able to capture the recruitment deviations estimated in the 2021 stock assessment.

In years without high recruitment estimates between 2011 and 2017, the catch plus sea-level hindcast had smaller standard deviations around the $\log_e$ bias-corrected recruitment deviations than the catch-only hindcast. However, in years with high recruitment estimates between 2011 and 2017, and for 2018 and 2019, the standard deviations around the $\log_e$ bias-corrected recruitment deviations were larger than those from the catch-only hindcast. The uncertainty in recruitment deviations from the 2021 models, and therefore in recruitment estimates, was larger in 2019 and 2020 due to the lack of fishery-independent survey data in 2020 and reduced survey effort in 2019 (Table S3, Fig. 7). Percent mean absolute relative differences from the catch plus sea-level hindcasts for recruitment, spawning stock biomass, and fraction of the unfished spawning biomass were, 33.1%, 43.3%, and 39.6%, respectively. Percent mean absolute relative differences from the catch-only hindcast for recruitment, spawning stock biomass, and fraction of the unfished spawning biomass were larger than those from the catch plus sea-level hindcast at 51.2%, 48.4%, and 40.2%, respectively. The lower percent mean absolute relative differences from the catch plus sea-level hindcasts indicate improved performance with respect to the 2021 stock assessment, in which these model derived estimates use all available data.

Discussion

A crux of fishery management is that while recruitment is seldom average, cohort strength is not well estimated until several years of data are available from surveys and fisheries. Thus, scientists and managers are always looking in the rear view mirror. The catch plus sea-level hindcast information presented here suggests that there is potential to improve forward-looking stock projections by better capturing stock trends, providing an improvement over the common practice of using the expected recruitment from a fitted stock-recruitment curve (average deterministic recruitment) when no other data are available to inform recruitment assumptions in stock projections.

Analyses of the relationships between sablefish and environmental drivers have generally focused on the northern portion of their West Coast range (Schirripa and Colbert 2006; Tolimieri et al. 2018; Haltuch et al. 2019b), either for *a priori* reasons (focusing on dynamics in the north because much of the age and length data come from the north) or because model fitting selected northern drivers. However, species distribution modeling of age-0 sablefish abundance using trawl survey data hints that dynamics south of Cape Mendocino are different and may also be important (Tolimieri et al. 2020). For 2003-2018, high coast-wide age-0 abundance was generally associated with high abundance north of Cape Mendocino. However, the northern-only models tend to over-predict recruitment in years of low abundance (e.g., 2005-2007) (Tolimieri et al. 2018; Haltuch et al. 2019b). These over-prediction years also had recruitment failures in the south suggesting that dynamics in the south may also be important but not adequately observed in the current data or captured in current modeling approaches. Our second-best candidate model did include DF4, or southern sea level. Recruitment failure in the south may be infrequent enough (Tolimieri et al. 2020) to limit the selection of southern drivers in model selection. Future modeling may look to evaluate processes in the south and integrate northern and southern predictors.

In other cases, the model under-predicted strong recruitments or over-predicted weak recruitments even though it did predict peaks or lows in those years. In addition to sea level and its consequences for larval dynamics, other biological mechanisms could provide additional predictive power for sablefish recruitment and stock size. For example, abundance of sablefish predators was generally low in 2006 and 2007, suggesting that we might expect good recruitment in these years (Haltuch et al. 2019b). However, the condition of age-7+ females was also low in these years (see Supplementary Material, Tables S4-S6, Figs. S8-S10 and Haltuch et al. 2019b). Note that adding same-year female condition as a predictor increased the model fit for 2003-2019 (r^2 = 0.44, see Supplementary Material) and resulted in better predictions for 2006 and 2007. It is not clear why female condition in late summer of the age-0 year would predict recruitment earlier in the year, but one hypothesis is that females were in poor enough condition earlier that they could not recover over the summer and that this poor condition resulted in lower egg production and potentially skip spawning (Rodgveller et al. 2016). It is also possible that the size- and age-structure of the spawning stock may play a role in recruitment dynamics, particularly if older or larger fish are more important to subsequent recruitment (Barneche et al. 2018; Ottersen and Holt 2022). Work in Alaska suggests that overwintering success for age-0 fishes (to age-1) is an important factor determining year-class strength (Callahan et al. 2021), which may also be a factor here. However, the abundance of age-0 fishes is correlated with the assessment-based recruitment estimates (there is some circularity), suggesting that overwintering success may be less important in the California Current (Haltuch et al. 2019b; Tolimieri et al. 2020). Nevertheless, many models assume consistent egg or larval production from spawners, while in reality both will likely be variable.

The population dynamics of sablefish on the US West Coast may also be linked to those of sablefish populations in Canada and Alaska, suggesting that additional factors beyond the northern sea-level index could improve on the analyses conducted here (Fenske et al. 2019). Sablefish recruitment on the West Coast, and in British Columbia, and Alaska exhibit some synchrony (Fenske et al. 2019; Goethel et al. 2020). For example, all three regions showed recruitment pulses in 2000 and 2008, but there are also lags in timing. Assessment models estimated strong year classes on the West Coast in 2013 and 2016, in British Columbia in 2013 and 2015, and in Alaska in 2014, 2016 and 2017 (Fig. S1). This variation in the timing of recent recruitment peaks may represent differences among regions in the timing of environmental conditions favorable to recruitment, but may also be artefacts of varying stock assessment modeling parameterizations across regions (Goethel et al. 2020). The oceanography related to strong sablefish recruitment does vary among regions (Shotwell et al. 2014; Coffin and Mueter 2015; Tolimieri et al. 2018), so an uncoupling of recruitment dynamics in the two regions is possible. Nevertheless, the general similarity in recruitment trends seen in Alaska, British Columbia, and the West Coast (Goethel et al. 2020) suggests that we need to be better understand connections in sablefish productivity across regions.

The inclusion of environmental drivers in stock assessment models has the potential to enhance the performance of these tools, which normally rely on a stock-recruitment relationship that does not vary with environmental variability (du Pontavice et al. 2022). Additionally, efforts to include environmental effects in stock assessments could benefit by including the environmental data analyses directly into the stock assessment. Another successful example that includes climate effects on recruitment is the improvement in predictions of recruitment and stock biomass for yellowtail flounder *Limanda ferruginea* in waters off of the northeastern USA due to incorporation of Cold Pool relationships (du Pontavice et al. 2022). In our work here, the catch plus sea-level hindcast was able to capture the increase in stock biomass from 2016 onward seen in the full 2021 sablefish assessment, while the catch-only hindcast predicted continued decline over the same period. The latter finding might erroneously imply the need for more conservative management of sablefish harvest. Including sea level also resulted in lower uncertainty for some assessment model parameters. These retrospective investigations provide a step towards understanding how climate data can inform stock projections for fishery management, and for general acceptance in moving from research to application. Furthermore, the Pacific Fisheries Management Council routinely uses catch-only projections to provide updated management advice between stock assessments; these catch-only updates rely on average recruitment assumptions. This work shows that environment-based indices of recruitment have the potential to provide fishery managers with improved leading information regarding incoming year class strength for informing decision making between stock assessments, thus bringing the management system closer to fishing targets. This work provides an example of how transitioning research products from research to operations can improve stock assessments and advice for fishery managers. The co-development of the science products and the management and decision-making frameworks that will use these scientific products and advice illustrate the benefits of frequent communication between fisheries scientists and fishery management bodies as we move towards climate-ready fisheries.

Acknowledgements

We thank 2 anonymous reviewers, J. Field, O. Hamel, C. Harvey, J. Samhouri, and S. Busch for comments on the manuscript. Thanks to E. Ward for help with running the dynamic factor analysis. Special thanks to G. Monmouth, C. de Troyes, R. de Boron, W. von Eschenbach, T.G. Poet, and T. Malory.

Competing interests

The authors declare there are not competing interests

Author roles

NT: Conceptualization, Formal Analysis, Writing-original draft

MH: Conceptualization, Formal Analysis, Writing-original draft

Funding statement

The authors declare no specific funding for this work.

Data availability statement

Sea-level data were obtained from NOAA Tides and Currents: NOAA Tides and Currents https://tidesandcurrents.noaa.gov/sltrends/

Sablefish data were obtained from the 2021 sablefish stock assessment available through the Pacific Fisheries Management Council: https://www.pcouncil.org/

Condition data in the Supplement were derived from the West Coast Groundfish Bottom Trawl Survey. Raw data are available via API from the FRAMD Data Warehouse: https://www.webapps.nwfsc.noaa.gov/data/map

References

Allen, M.J., and Smith, G.B. 1988. Atlas and zoogeography of common fishes in the Bering Sea and northeastern Pacific. NOAA Tech Rep 66, National Marine Fisheries Service Seattle, WA.

Barneche, D.R., Robertson, D.R., White, C.R., and Marshall, D.J. 2018. Fish reproductive-energy output increases disproportionately with body size. Science **360**(6389): 642-645. doi:10.1126/science.aao6868.

Boehlert, G.W., and Yoklavich, M.M. 1985. Larval and juvenile growth of sablefish, *Anoplopoma fimbria*, as determined from otolith increments. Fish. Bull. **83**(3): 475-481.

Bradburn, M.J., Keller, A., and Horness, B.H. 2011. The 2003 to 2008 U.S. West Coast bottom trawl surveys of groundfish resources off Washington, Oregon, and California: Estimates of distribution, abundance, length, and age composition. U.S. Deptartment of Commerce.

Brock, V.E. 1940. Note on the young of the sablefish, *Anoplopoma fimbria*. Copea **1940**: 268-270. doi:10.2307/1438590.

Callahan, M.W., Beaudreau, A.H., Heintz, R., and Mueter, F. 2021. First winter energy allocation in juvenile sablefish *Anoplopoma fimbria*, a fast growing marine piscivore. Mar. Ecol. Prog. Ser. **663**: 145-156. doi:10.3354/meps13641.

Chelton, D.B. 1984. Seasonal variability of alongshore geostrophic velocity off central California. J. Geophys. Res-Oceans **89**(C3): 3473-3486. doi:10.1029/JC089iC03p03473.

Chelton, D.B., and Davis, R.E. 1982. Monthly mean sea-level variability along the west coast of North America. J. Phys. Oceanogr. **12**(8): 757-784. doi:10.1175/1520-0485(1982)012<0757:Mmslva>2.0.Co;2.

Chelton, D.B., Bernal, P.A., and Mcgowan, J.A. 1982. Large-scale interannual physical and biological interaction in the California Current. J. Mar. Res. **40**(4): 1095-1125.

Coffin, B., and Mueter, F. 2015. Environmental covariates of sablefish (*Anoplopoma fimbria*) and Pacific ocean perch (*Sebastes alutus*) recruitment in the Gulf of Alaska. Deep Sea Res. II **132**: 194-209. doi:10.1016/j.dsr2.2015.02.016i.

Connolly, T.P., Hickey, B.M., Shulman, I., and Thomson, R.E. 2014. Coastal trapped waves, alongshore pressure gradients, and the California Undercurrent. J. Phys. Oceanogr. **44**(1): 319-342. doi:10.1175/Jpo-D-13-095.1.

du Pontavice, H., Miller, T.J., Stock, B.C., Chen, Z., and Saba, V.S. 2022. Ocean model-based covariates improve a marine fish stock assessment when observations are limited. ICES J. Mar. Sci. **79**(4): 1259-1273. doi:10.1093/icesjms/fsac050.

Fenske, K.H., Berger, A.M., Connors, B., Cope, J.M., Cox, S.P., Haltuch, M.A., Hanselman, D.H., Kapur, M., Lacko, L., Lunsford, C., Rodgveller, C., and Williams, B. 2019. Report on the 2018 International Sablefish Workshop. NOAA Tech. Mem. NMFS AFSC **387**: p. 107.

Goethel, D.R., Hanselman, D.H., Rodgveller, C., Fenske, K.H., Shotwell, S.K., Echave, K.B., Malecha, P.W., Siwicke, K.A., and Lunsford, C.R. 2020. Assessment of the sablefish stock in Alaska. Pages 1-257 NPFMC Bering Sea, Aleutian Islands and Gulf of Alaska SAFE. Natinal Marine Fisheries Service, Alaska Fisheries Science Center.

Grover, J.J., and Olla, B.L. 1987. Effects of an El Nino event on the food habits of larval Sablefish, *Anoplopoma fimbria*, off Oregon and Washington. Fish. Bull. **85**(1): 71-79.

Haltuch, M.A., and Punt, A.E. 2011. The promises and pitfalls of including decadal-scale climate forcing of recruitment in groundfish stock assessment. Can. J. Fish. Aquat. Sci. **68**(5): 912-926. doi:10.1139/F2011-030.

Haltuch, M.A., Tolimieri, N., Lee, Q., and Jacox, M.G. 2019a. Oceanographic drivers of petrale sole recruitment in the California Current Ecosystem. Fish. Oceanogr. **29**(2): 122-136. doi:10.1111/fog.12459.

Haltuch, M.A., Johnson, K.F., Tolimieri, N., Kapur, M.S., and Castillo-Jordán, C.A. 2019b. Status of the sablefish stock in U.S. waters in 2019. Pacific Fisheries Management Council, Portland, OR, 7700 Ambassador Place NE, Suite 200.

Hart, J.L. 1973. Pacific fishes of Canada. Fisheries Research Board of Canada, Bulletin 180, St. Andrews, New Brunswick, Canada.

Head, M.A., Keller, A.A., and Bradburn, M. 2014. Maturity and growth of sablefish, *Anoplopoma fimbria*, along the US West Coast. Fish. Res. **159**: 56-67. doi:10.1016/j.fishres.2014.05.007.

Holmes, E.E., Ward, E.J., and Scheuerell, M.D. 2021. Analysis of multivariate time-series using the MARSS package, version 3.11.4. NOAA Fisheries, Northwest Fisheries Science Center, 2725 Montlake Blvd E., Seattle, WA 98112.

Hunter, J.R., Macewicz, B.J., and Kimbrell, C.A. 1989. Fecundity and other aspects of the reproduction of sablefish, *Anoplopoma fimbria,* in Central California waters. Cal. Coop. Ocean. Fish. **30**: 61-72.

Jasonowicz, A.J., Goetz, F.W., Goetz, G.W., and Nichols, K.M. 2017. Love the one you're with: genomic evidence of panmixia in the sablefish (*Anoplopoma fimbria*). Can. J. Fish. Aquat. Sci. **74**(3): 377-387. doi:10.1139/cjfas-2016-0012.

Johnson, K.F., Rudd, M.B., Pons, M., Akselrud, C.A., Lee, Q., Hurtado-Ferro, F., Haltuch, M.A., and Hamel, O.S. 2016. Status of the U.S. sablefish resource in 2015. Pacific Fisheries Management Council, Portland, OR.

Kapur, M., Haltuch, M., Connors, B., Rogers, L., Berger, A., Koontz, E., Cope, J., Echave, K., Fenske, K., Hanselman, D., and Punt, A.E. 2020. Oceanographic features delineate growth zonation in Northeast Pacific sablefish. Fish. Res. **222**. doi:10.1016/j.fishres.2019.105414.

Kapur, M.S., Lee, Q., Correa, G.M., Haltuch, M., Gertseva, V., and Hamel, O.S. 2021. Status of Sablefish (*Anoplopoma fimbria*) along the US West coast in 2021. Pacific Fisheries Management Council, Portland, Oregon.

Keister, J.E., Di Lorenzo, E., Morgan, C.A., Combes, V., and Peterson, W.T. 2011. Zooplankton species composition is linked to ocean transport in the Northern California Current. Global Change Biol. **17**(7): 2498-2511. doi:10.1111/j.1365-2486.2010.02383.x.

Keller, A.A., Wallace, J.R., and Methot, R.D. 2017. The Northwest Fisheries Science Center’s West coast groundfish bottom trawl survey: History design, and description. U.S. Department of Commerce.

Kendall, A.W., and Matarese, A.C. 1987. Biology of eggs, larvae, and epipelagic juveniles of sablefish, *Anoplopoma fimbria*, in relation to their potential use in management. Mar. Fish. Rev. **49**(1): 1-13.

King, J.R., McFarlane, G.A., and Beamish, R.J. 2000. Decadal-scale patterns in the relative year class success of sablefish (*Anoplopoma fimbria*). Fish. Oceanogr. **9**(1): 62-70.

Mason, J.C., Beamish, R.J., and Mcfarlane, G.A. 1983. Sexual maturity, fecundity, spawning, and early life-history of sablefish (*Anoplopoma fimbria*) off the Pacific Coast of Canada. Can. J. Fish. Aquat. Sci. **40**(12): 2126-2134. doi:10.1139/f83-247.

Mcfarlane, G.A., and Beamish, R.J. 1992. Climatic influence linking copepod production with strong year-classes in sablefish, *Anoplopoma fimbria*. Can. J. Fish. Aquat. Sci. **49**(4): 743-753. doi:10.1139/f92-083.

Methot, R.D., and Taylor, I.G. 2011. Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Can. J. Fish. Aquat. Sci. **68**(10): 1744-1760. doi:10.1139/F2011-092.

Methot, R.D., and Wetzel, C.R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fish. Res. **142**: 86-99. doi:10.1016/j.fishres.2012.10.012.

Mitchell, C.T., and Hunter, J.R. 1970. Fishes associated with drifting kelp, *Macrocystis pyfifera*, off the coast of southern California and northern Baja California. Calif. Fish Game **56**: 288-297.

Moser, H.G., Charter, R.L., Smith, P.E., Lo, N.C.H., Ambrose, D.A., Meyer, C.A., Sandknop, E.M., and Watson, W. 1994. Early life history of sablefish, *Anoplopoma fimbria*, off Washington, Oregon, and California, with application to biomass estimation. Cal. Coop. Ocean. Fish. **35**: 144-159.

Neveu, E., Moore, A.M., Edwards, C.A., Fiechter, J., Drake, P., Crawford, W.J., Jacox, M.G., and Nuss, E. 2016. An historical analysis of the California Current circulation using ROMS 4D-Var: System configuration and diagnostics. Ocean Model. Online **99**: 133-151. doi:10.1016/j.ocemod.2015.11.012.

Ottersen, G., and Holt, R.E. 2022. Long-term variability in spawning stock age structure influences climate–recruitment link for Barents Sea cod. Fish. Oceanogr. **32**(1): 91-105. doi:10.1111/fog.12605.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ralston, S., Sakuma, K.M., and Field, J.C. 2013. Interannual variation in pelagic juvenile rockfish (*Sebastes* spp.) abundance - going with the flow. Fish. Oceanogr. **22**(4): 288-308. doi:10.1111/fog.12022.

Reid, J.L., and Mantyla, A.W. 1976. The effect of the geostrophic flow upon coastal sea elevations in the northern North Pacific Ocean. J. Geophys. Res. **81**(18): 3100-3110. doi:10.1029/JC081i018p03100.

Rodgveller, C.J., Stark, J.W., Echave, K.B., and Hulson, P.J.F. 2016. Age at maturity, skipped spawning, and fecundity of female sablefish (*Anoplopoma fimbria*) during the spawning season. Fish. Bull. **114**(1): 89-102. doi:10.7755/Fb.114.1.8.

Schirripa, M.J. 2007. Status of the sablefish resource off the Continental U.S. Pacific coast in 2007. NOAA Fisheries, Northwest Fisheries Science Center, Newport, OR.

Schirripa, M.J., and Colbert, J.J. 2005. Status of the sablefish resource off the continental U.S. Pacific Coast in 2005. Pacific Fisheries Management Council.

Schirripa, M.J., and Colbert, J.J. 2006. Interannual changes in sablefish (*Anoplopoma fimbria*) recruitment in relation to oceanographic conditions within the California Current System. Fish. Oceanogr. **15**(1): 25-36. doi:10.1111/j.1365-2419.2005.00352.x.

Schirripa, M.J., Goodyear, C.P., and Methot, R.M. 2009. Testing different methods of incorporating climate data into the assessment of US West Coast sablefish. ICES J. Mar. Sci. **66**(7): 1605-1613. doi:10.1093/icesjms/fsp043.

Schroeder, I.D., Santora, J.A., Bograd, S.J., Hazen, E.L., Sakuma, K.M., Moore, A.M., Edwards, C.A., Wells, B.K., and Field, J.C. 2019. Source water variability as a driver of rockfish recruitment in the California Current Ecosystem: implications for climate change and fisheries management. Can. J. Fish. Aquat. Sci. **76**(6): 950-960. doi:10.1139/cjfas-2017-0480.

Shotwell, S.K., Hanselman, D.H., and Belkin, I.M. 2014. Toward biophysical synergy: Investigating advection along the Polar Front to identify factors influencing Alaska sablefish recruitment. Deep Sea Research II **107**: 40-53. doi:10.1016/j.dsr2.2012.08.024.

Stewart, I.J., Thorson, J.T., and Wetzel, C. 2011. Status of the U.S. sablefish resource in 2011. Pacific Fisheries Management Council, Portland, OR.

Stige, L.C., Hunsicker, M.E., Bailey, K.M., Yaragina, N.A., and Hunt, G.L. 2013. Predicting fish recruitment from juvenile abundance and environmental indices. Mar. Ecol. Prog. Ser. **480**: 245-261. doi:10.3354/meps10246.

Tolimieri, N. 2006. Patterns in species richness, species density, and evenness in groundfish assemblages on the continental slope of the U.S. Pacific coast. Environ. Biol. Fishes **78**(3): 241-256. doi:10.1007/s10641-006-9093-5.

Tolimieri, N., and Levin, P.S. 2006. Assemblage structure of eastern Pacific groundfishes on the U.S. continental Slope in Relation to physical and environmental variables. Trans. Am. Fish. Soc. **135**(2): 317-332. doi:10.1577/t05-092.1.

Tolimieri, N., Wallace, J., and Haltuch, M. 2020. Spatio-temporal patterns in juvenile habitat for 13 groundfishes in the California Current Ecosystem. Plos One **15**(8). doi:10.1371/journal.pone.0237996.

Tolimieri, N., Haltuch, M.A., Lee, Q., Jacox, M.G., and Bograd, S.J. 2018. Oceanographic drivers of sablefish recruitment in the California Current. Fish. Oceanogr. **27**(5): 458-474. doi:10.1111/fog.12266.

Ward, E.J., Anderson, S.C., Damiano, L.A., and Malick, M.J. 2021. bayesdfa: Bayesian Dynamic Factor Analysis (DFA) with 'Stan'. R package version 1.2.0. https://CRAN.R-project.org/package=bayesdfa.

Zuur, A.F., Tuck, I.D., and Bailey, N. 2003a. Dynamic factor analysis to estimate common trends in fisheries time series. Can. J. Fish. Aquat. Sci. **60**(5): 542-552. doi:10.1139/F03-030.

Zuur, A.F., Fryer, R.J., Jolliffe, I.T., Dekker, R., and Beukema, J.J. 2003b. Estimating common trends in multivariate time series using dynamic factor analysis. Environmetrics **14**(7): 665-685. doi:10.1002/env.611.

Tables

Table 1. Coefficients for the best-fit model including bias estimates.

Predictor	Coefficient	Bias	SE
Intercept	0.240	-0.004	0.163
DF1	-0.642	-0.027	0.191

Figure Captions

Fig. 1. Time series of sablefish a) spawning biomass, b) age-0 abundance, and c) recruitment deviations, and d) the relationship between spawning biomass and age-0 abundance. Data are from Table 18 in 2021 sablefish stock assessment (Kapur et al. 2021).

Fig. 2. Location of tide gauges used in the sea-level analyses. Map was prepared using R software (R Core Team 2021) and the ‘maps’ package using WGS84 datum and a rectangular projection with longitude and latitude scales are equivalent at the center of the picture.

Fig. 3. Dynamic factors for the best-fit DFA model reducing 16 sea-level time series to five common trends. Because the data were normalized prior to analyses, the displayed data are dimensionless, scaled trends.

Fig. 4. Loadings for the five dynamic factors for the best-fit DFA model reducing 16 sea-level time series to five common trends.

Fig. 5. Performance of best-fit model (r^2 = 0.15). Solid black line is the predicted recruitment deviations from the best-fit model with 95% confidence limits.

Fig. 6. Relationship between the first dynamic factor summarizing variation in sea level in the north and sablefish recruitment deviations (r^2 = 0.15).

Fig. 7. Panel (a) shows the time series of the fraction of unfished biomass estimated from four model runs based off the 2021 stock assessment, (b) shows log recruitment deviations estimated from the same models. Note, the black and grey lines/points overlap substantially in both panels.

Figures

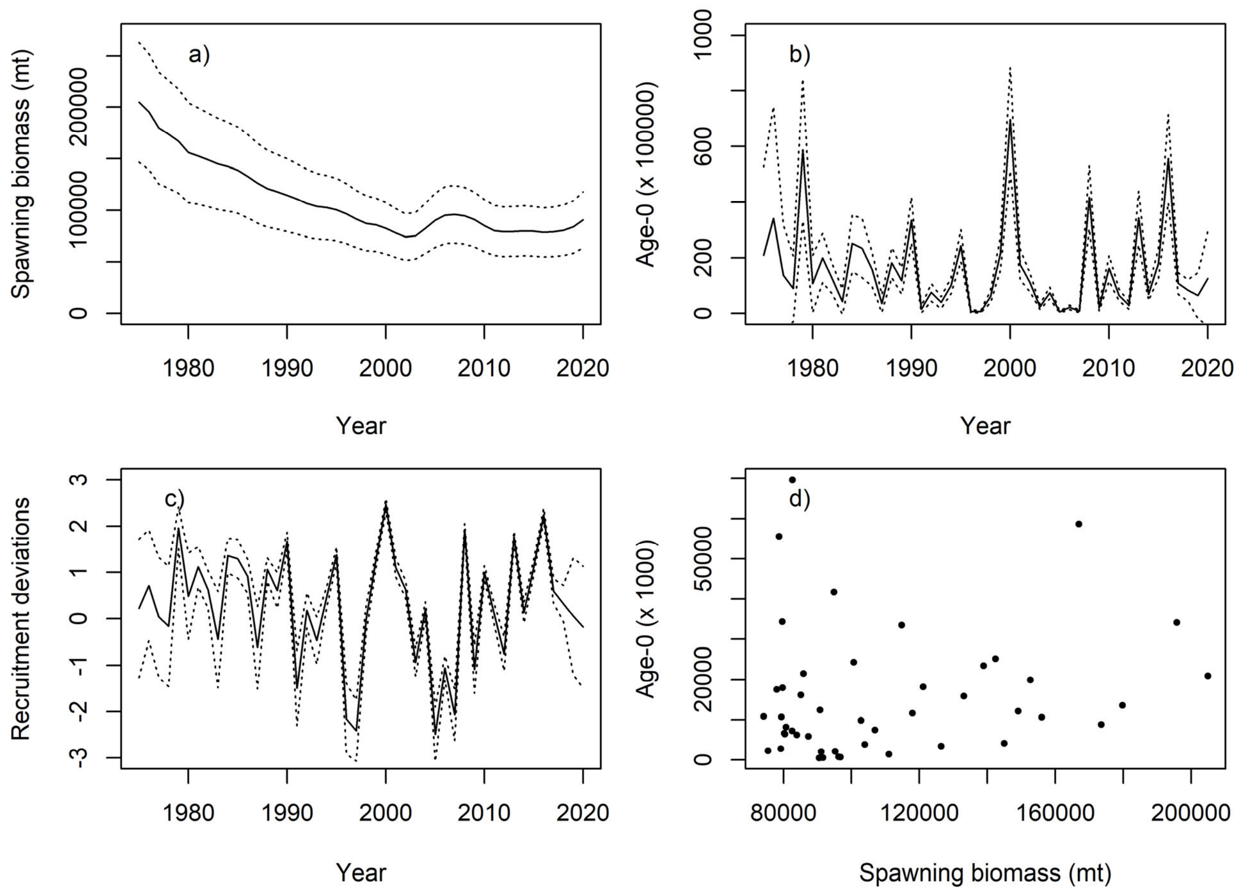

Figure 1

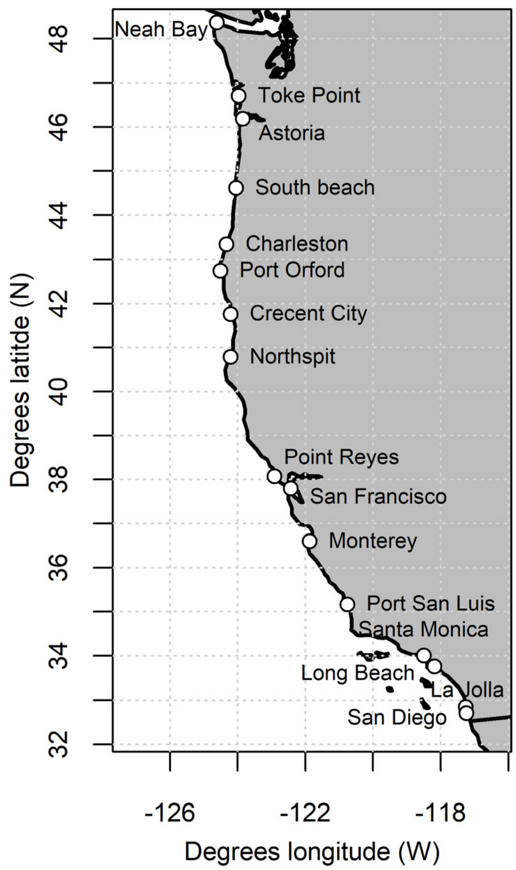
Neah Bay
Toke Point
Astoria
South beach
Charleston
Port Orford
Crecent City
Northspit
Point Reyes
San Francisco
Monterey
Port San Luis
Santa Monica
Long Beach
La Jolla
San Diego
Degrees latitude (N)
48
46
44
42
40
38
36
34
32
-126
-122
-118
Degrees longitude (W)

Figure 2

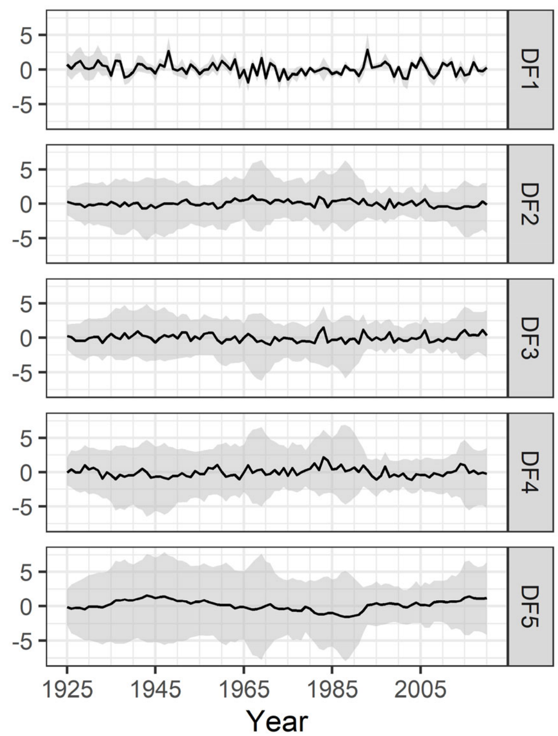

Figure 3

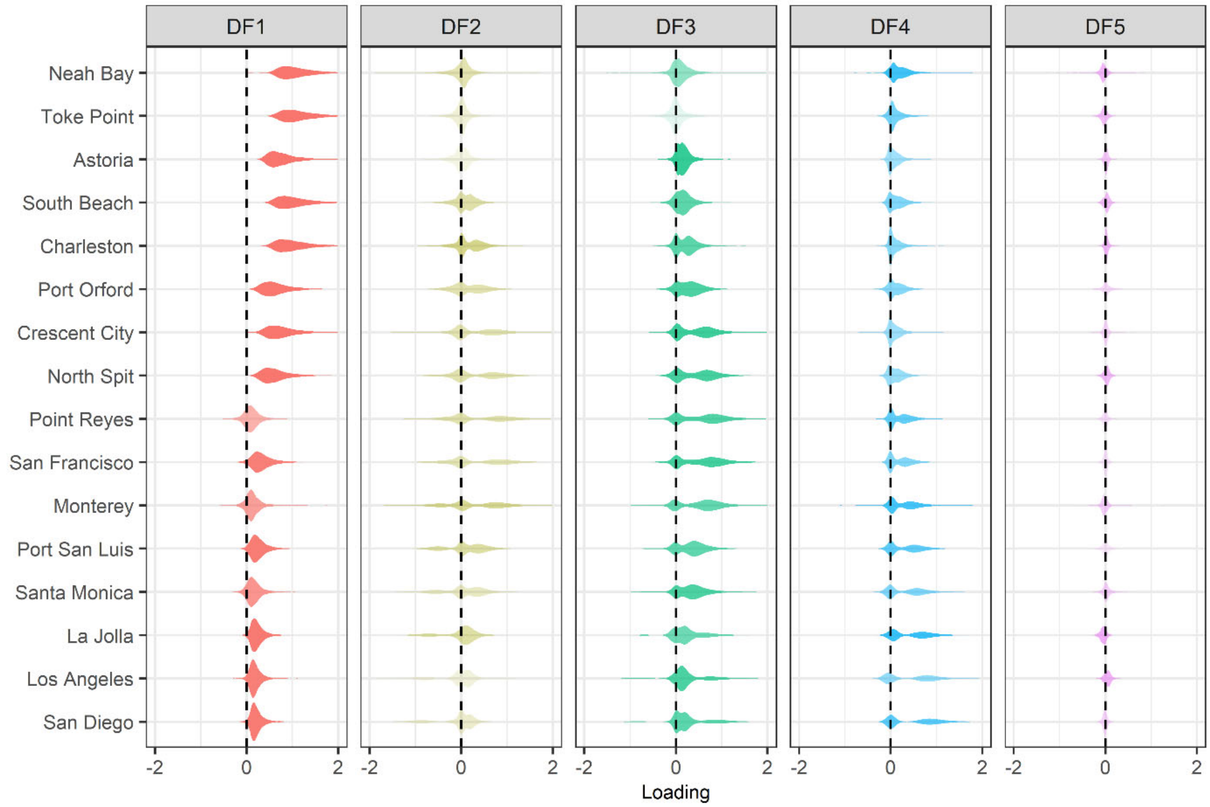
DF1
DF2
DF3
DF4
DF5
Neah Bay
Toke Point
Astoria
South Beach
Charleston
Port Orford
Crescent City
North Spit
Point Reyes
San Francisco
Monterey
Port San Luis
Santa Monica
La Jolla
Los Angeles
San Diego
-2
0
2
Loading

Figure 4

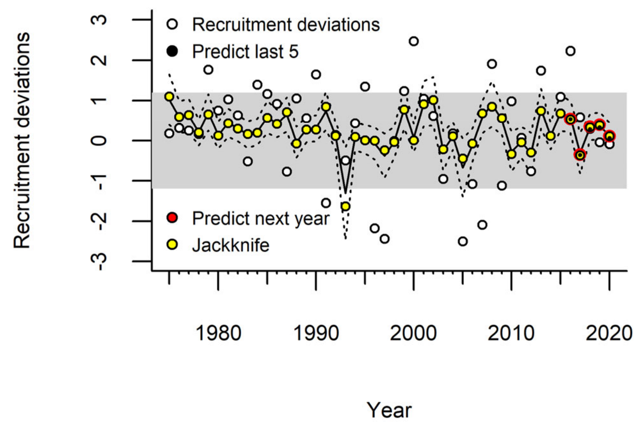

Figure 5

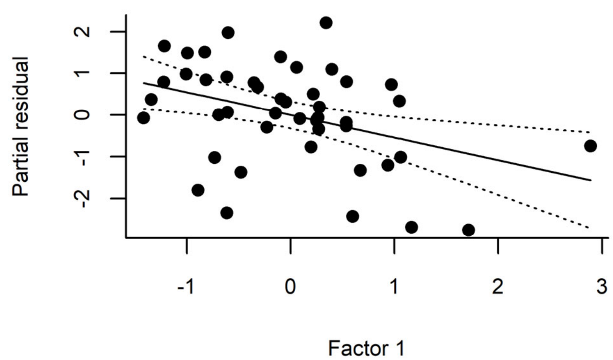

Figure 6

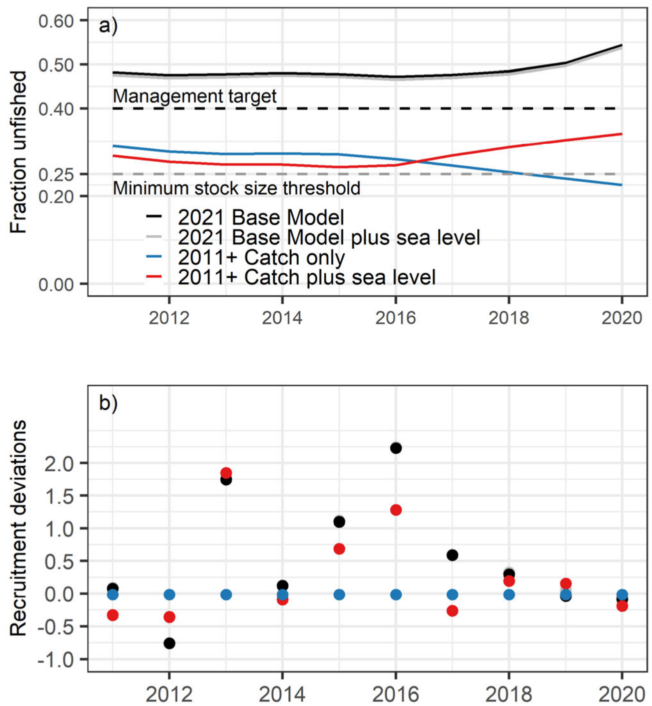
a)
0.60
0.50
0.40
0.25
0.20
0.00
Fraction unfished
Management target
Minimum stock size threshold
2021 Base Model
2021 Base Model plus sea level
2011+ Catch only
2011+ Catch plus sea level
2012
2014
2016
2018
2020
b)
Recruitment deviations
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
2012
2014
2016
2018
2020

Figure 7

Supplementary Material

Synchrony among regions

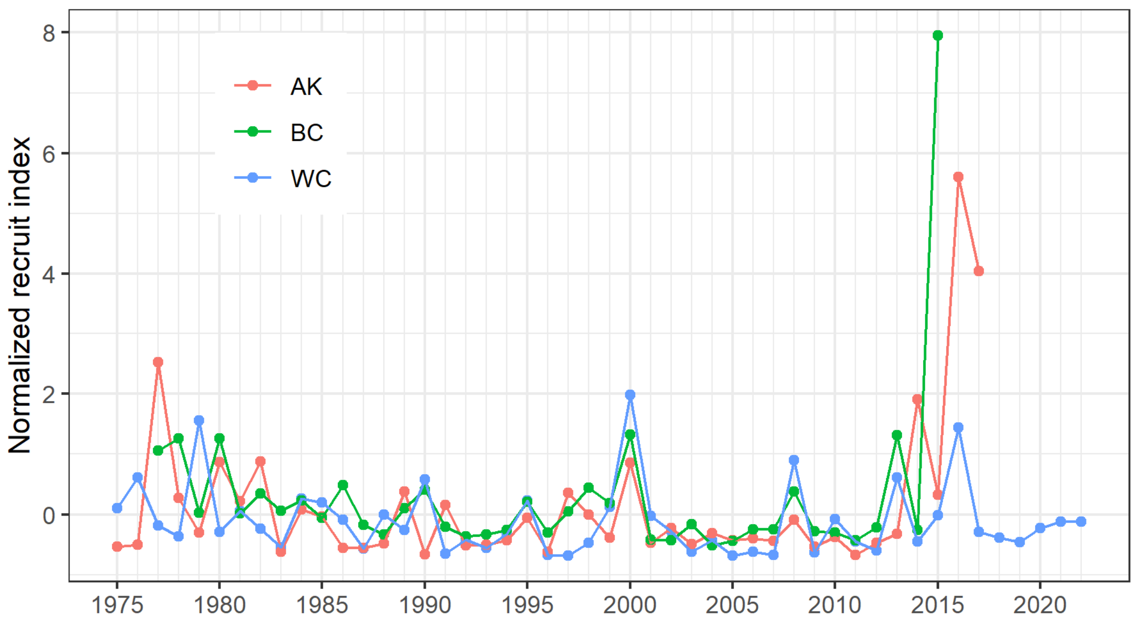

Fig. S1. Normalized recruitment indices from the Alaskan, British Columbia, and West Coast stock assessments. Data from Goethel et al. (2020) & Kapur et al. (2021).

Sea level time series

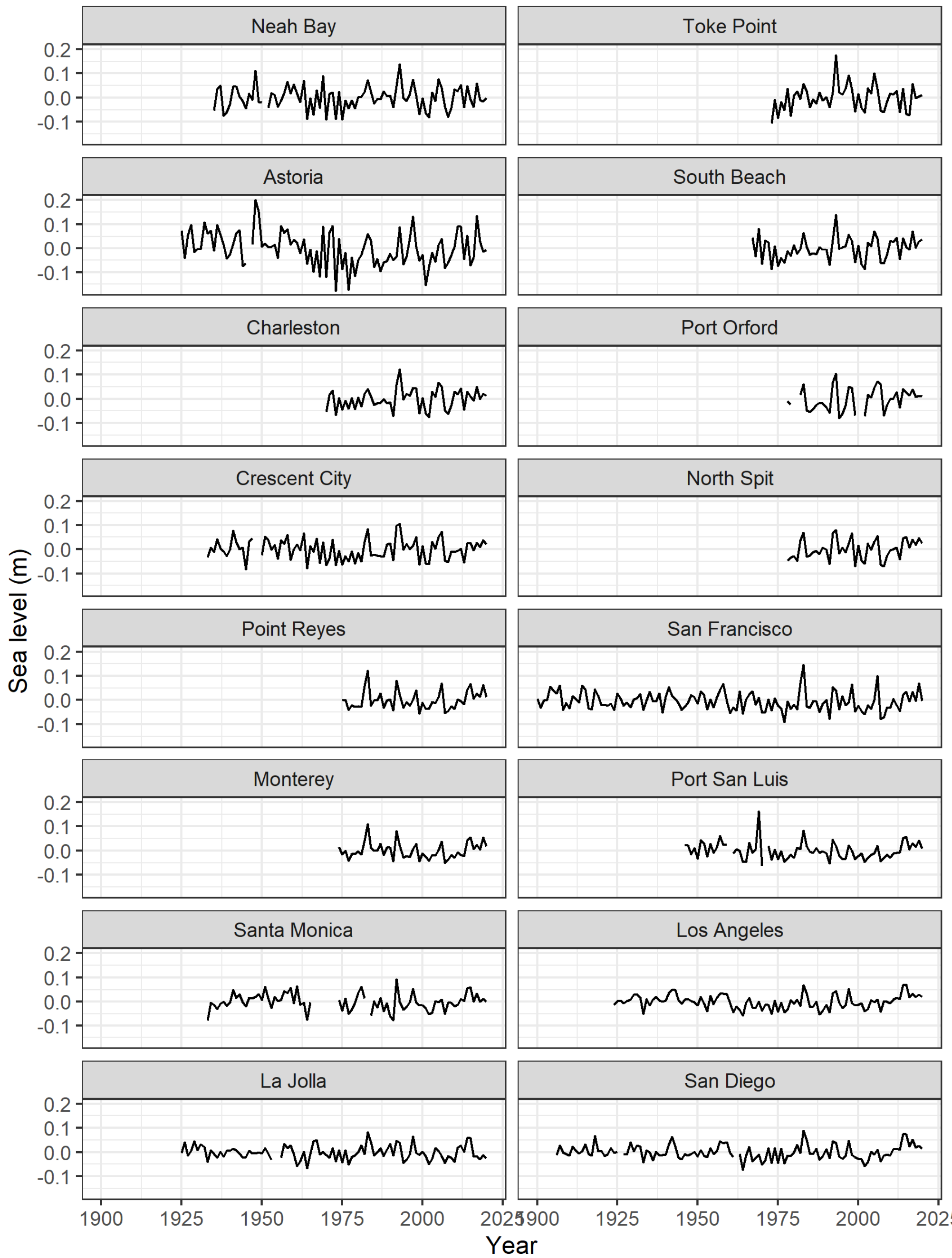

Fig. S2. Mean monthly sea level in the second quarter (April-June) at 16 stations along the US west coast from 1900 to 2019. Average seasonal cycle and linear trend have been removed.

Model fit to the data

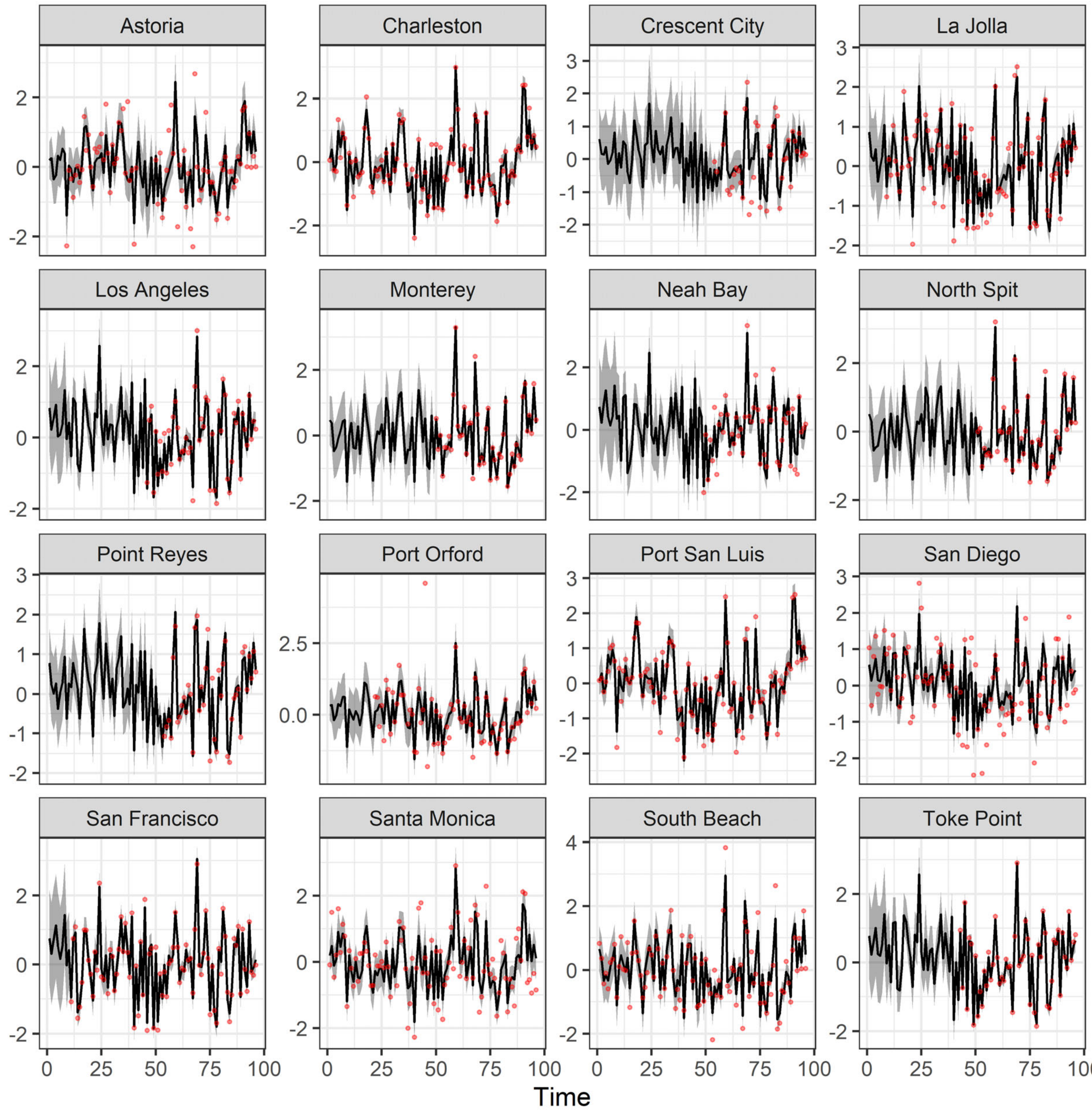

Fig. S3. Fit of the DFA model (black line) to the observed data (red points) for 16 tide gauge locations along the West Coast of the U.S.A.

Model testing and validation

We ran an array of additional tests to validate the best-fit model (Model 1). Some model validation actions are described in the main text. Here, we describe three additional validation tests:

The recruitment deviations used in the main analysis were assessment-based estimates and have error. Therefore, we determined whether the precision of recruitment deviations from the assessment model affected the recruitment-environment relationship. We resampled recruitment deviations from a normal distribution for each year using the recruitment deviation and its standard error from the 2019 assessment. We then refit the model 1000 times and compared the r^2 values. Median r^2 was $r^2 = 0.16$ ($CI_{95\%} = 0.6 – 0.28$).

To determine whether individual years had a strong influence on which terms (DFs) were included in the best-fit model, we jackknifed years and re-ran the entire model selection exercise 1000 times for each of the 45 years. We then compared what terms were included in the model from each iteration that had the lowest AICc. The first dynamic factor (DF1) was included in all 45 models. DF2 and DF4 each occurred in one model each.

Finally, we combined the two preceding analyses. We re-ran the entire model fitting exercise 1000 times using the re-sampled sablefish recruitment deviations. We then compared the best-fit (in this case lowest AICc) models from each run and determined the number of times each DF appeared in the model with the lowest AICc. DF1 was included in over 95% of all best-fit models, while other terms appear more sporadically. Note these results are for the model from each iteration with the lowest AICc not the lowest AICc and fewest parameters. When the best-fit model was chosen based on delta AICc < 2.0 and the fewest parameters, over 90% of models contained only DF1 (Table S1).

Table S1. Results of jackknife-refit analysis showing the number of times the predictor was included in the best-fit model (lowest AICc) out of 1000 iterations.

Predictor	Number of models lowest AICc	Number of models lowest AICc & fewest parameters
DF1	957	906
DF12	30	2
DF2	217	39
DF22	5	0
DF3	23	5
Df32	12	2
DF4	243	54
DF42	115	16
DF5	191	50
DF5	51	16

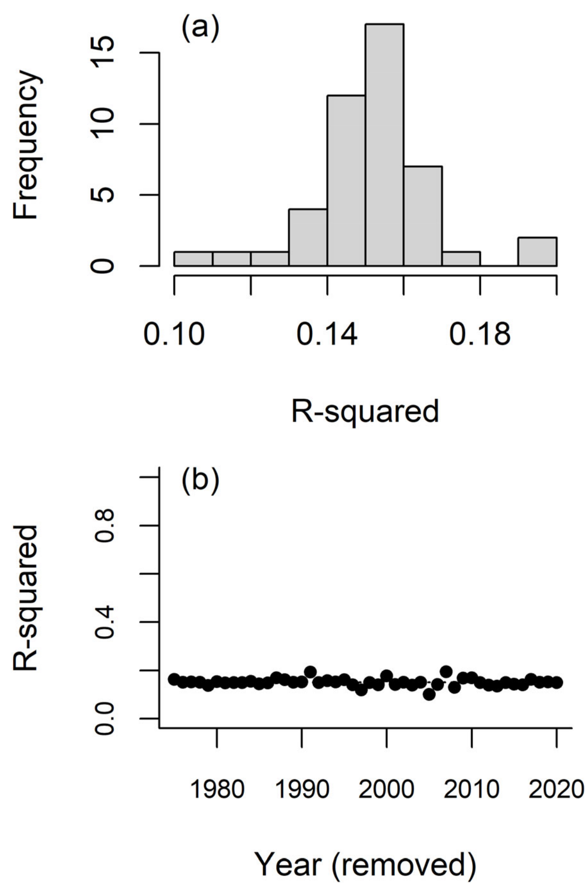

Fig. S4. Results of jackknife refitting of the best-fit model for 1975-2020.

Standard model diagnostics for the sea level recruitment model

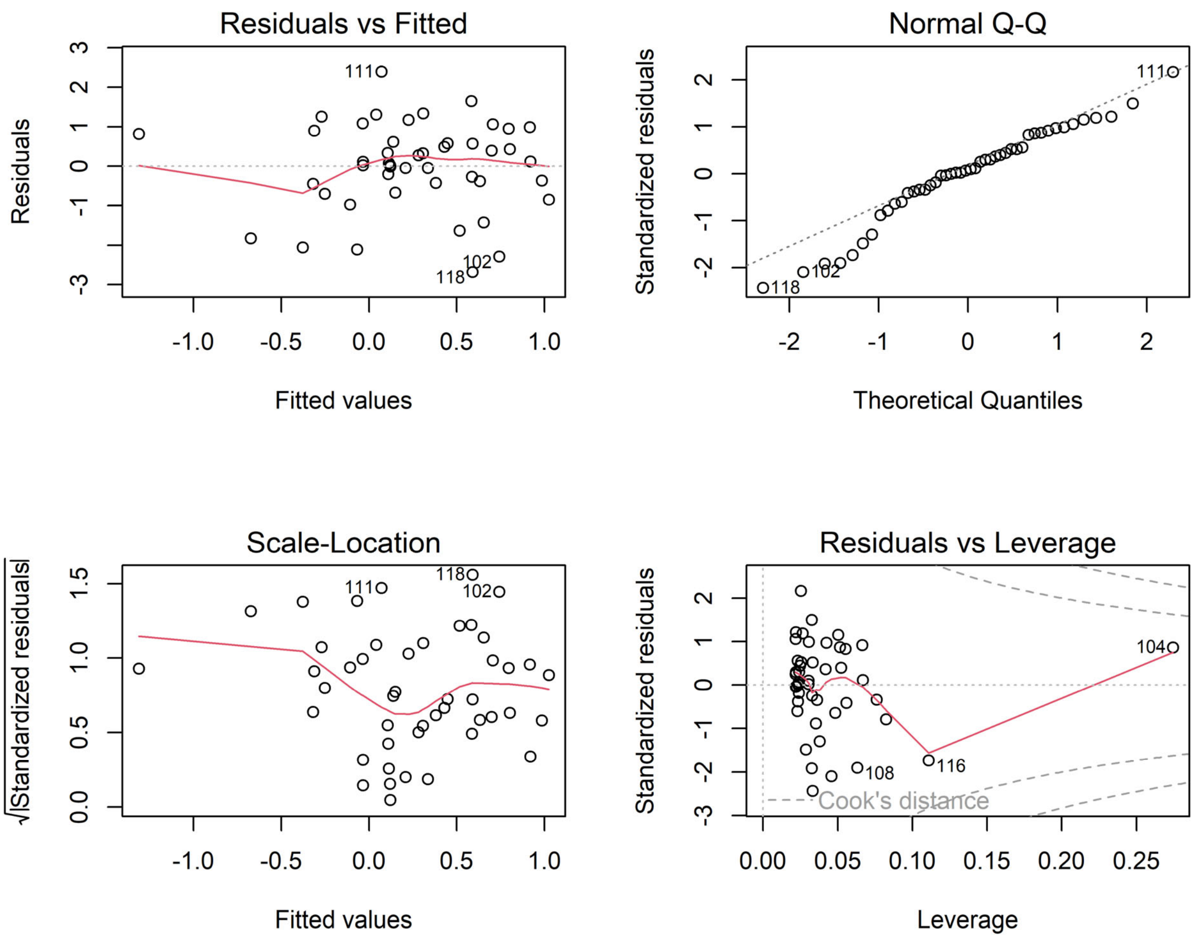

Fig. S5. Plots of model diagnostics for the best-fit model: recruitment deviations = DF1.

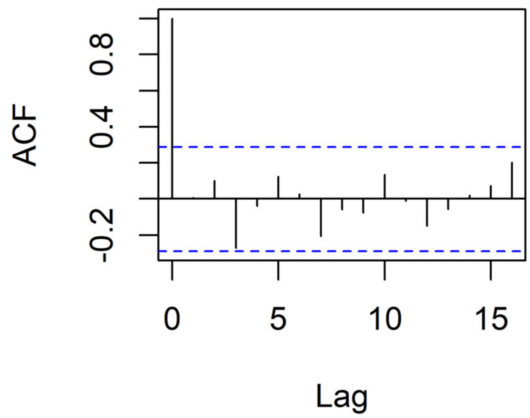

Fig. S6. Autocorrelations factors for the best-fit model.

Stock assessment output

Removing the sea-level index from the 2021 assessment had little impact on model estimates of natural mortality and growth parameters, but did suggest slightly lower recruitments during 2011-2019 (Table S2). In 2020, when there were no survey data, the model with the sea-level index showed slightly lower recruitment than the model without (Table S3). As the time series of fishery-independent and -dependent data available to the model decreased, model estimates of natural mortality increased, while estimates for the Von Bertalanffy k parameters increased (Table 2), resulting in decreases in estimated unfished spawning biomass and stock status (Table S3). The standard deviations for natural mortality and growth parameters generally increased as the time series of available data declined (Table S2).

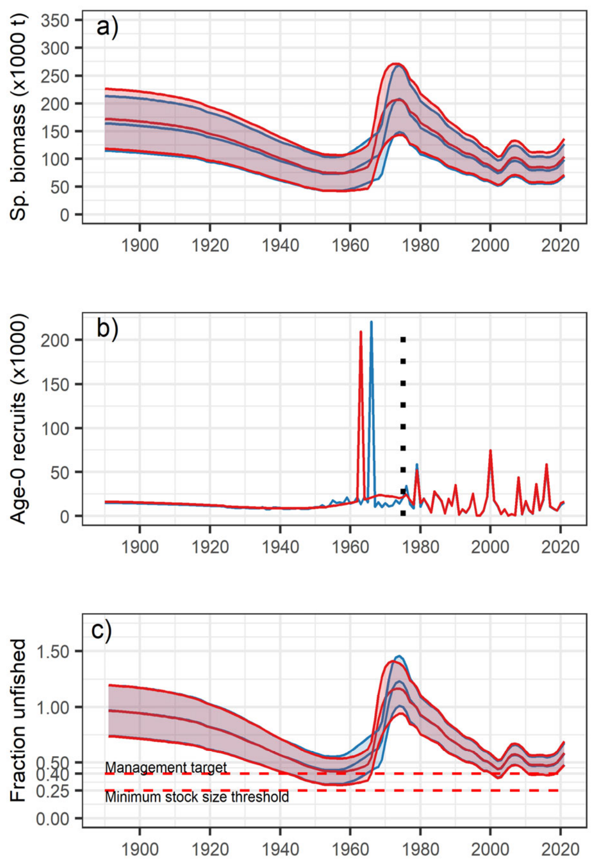

Fig. S7. Comparisons of the time series of spawning biomass (top row), age-0 recruits (middle row), and stock depletion (bottom row) between the 2021 stock assessment model used for management advice that includes sea level (blue lines) and a model sensitivity run with the sea-level index removed (red lines) (Kapur et al. 2021). Dotted black line in (b) indicates first year of recruitment deviations used in the analyses.

Table S2. Select parameter estimates from the 2021 and 2011 stock assessment model runs. Bold values represent years with reduced survey effort (2019) and no survey (2020).

	2021 Assessment	2021 Assessment, No DF1 sea level	2011 Catch and DF1 sea level hindcast	2011 Catch only hindcast	2021 Assessment	2021 Assessment, No DF1 sea level	2011 Catch and DF1 sea level hindcast	2011 Catch only hindcast
	Estimates				Standard Deviations			
Female Natural Morality	0.073	0.072	0.098	0.095	0.008	0.008	0.010	0.010
Female growth at minimum age	25.7	25.7	24.9	24.9	0.456	0.455	0.515	0.515
Female growth at maximum age	62.5	62.5	63.2	63.2	0.633	0.633	0.625	0.625
Female VonBertlanffy k	0.343	0.343	0.380	0.380	0.015	0.015	0.016	0.016
Male Natural Morality	0.060	0.060	0.082	0.079	0.006	0.006	0.008	0.008
Male growth at minimum age	26.9	26.9	26.0	26.0	0.514	0.515	0.704	0.703
Male growth at maximum age	56.6	56.6	56.9	56.9	0.322	0.323	0.325	0.325
Male VonBertlanffy k	0.371	0.371	0.419	0.420	0.014	0.014	0.017	0.017
SR_LN(R0)	9.705	9.700	9.979	9.846	0.305	0.304	0.313	0.303
2011 Recruitment Deviation	0.09	0.09	-0.33	-0.019	0.221	0.221	1.196	1.393
2012 Recruitment Deviation	-0.76	-0.75	-0.36	-0.019	0.363	0.362	1.193	1.393
2013 Recruitment Deviation	1.76	1.76	1.85	-0.019	0.130	0.130	0.866	1.393
2014 Recruitment Deviation	0.13	0.13	-0.09	-0.019	0.226	0.226	1.322	1.393
2015 Recruitment Deviation	1.12	1.11	0.69	-0.019	0.167	0.167	1.972	1.393
2016 Recruitment Deviation	2.25	2.24	1.28	-0.019	0.137	0.138	1.171	1.393
2017 Recruitment Deviation	0.60	0.60	-0.27	-0.019	0.264	0.263	1.237	1.393

2018 Recruitment Deviation	0.32	0.31	0.19	-0.019	0.397	0.398	1.374	1.393
2019 Recruitment Deviation	**0.05**	**-0.04**	0.15	-0.019	**1.255**	**1.235**	1.440	1.393
2020 Recruitment Deviation	**-0.19**	**-0.10**	-0.19	-0.019	**1.316**	**1.392**	1.251	1.393

Table S3. Select derived estimates from the 2021 and 2011 stock assessment model runs. Bold values represent years with reduced survey effort (2019) and no survey (2020).

	2021 Assessment	2021 Assessment, No DF1 sea level	2011 Catch and DF1 sea level hindcast	2011 Catch only hindcast	2021 Assessment	2021 Assessment, No DF1 sea level	2011 Catch and DF1 sea level hindcast	2011 Catch only hindcast
	Estimates				Standard Deviations			
Unfished Spawning Biomass (mt)	168,875	168,484	158,521	145,676	31,187	30,956	31,820	28,653
Unfished Recruitment (mt)	16,392	16,316	21,571	18,889	5,003	4,956	6,758	5,726
Recruitment								
2011	6,446	6,427	4,951	6,042	2,147	2,130	6,133	8,698
2012	2,759	2,767	4,735	5,967	1,227	1,225	5,854	8,592
2013	34,308	33,934	42,799	5,934	9,685	9,521	38,892	8,546
2014	6,709	6,685	6,126	5,944	2,281	2,262	8,395	8,561
2015	18,011	17,774	13,334	5,929	5,450	5,351	27,181	8,544
2016	55,595	55,061	24,165	5,867	15,803	15,574	29,572	8,460
2017	10,689	10,689	5,277	5,775	3,906	3,885	6,786	8,336
2018	8,151	7,966	8,492	5,669	3,894	3,805	12,144	8,195
2019	**6,274**	**5,674**	8,282	5,560	**8,224**	**7,319**	12,410	8,050
2020	**12,455**	**13,563**	14,761	13,539	**17,074**	**19,633**	19,264	19,642
Fraction of unfished spawning biomass								
2011	0.476	0.473	0.292	0.315	0.081	0.081	0.064	0.071
2012	0.469	0.467	0.279	0.301	0.081	0.081	0.065	0.072

2013	0.471	0.469	0.272	0.296	0.082	0.081	0.067	0.074
2014	0.475	0.472	0.272	0.297	0.082	0.081	0.069	0.077
2015	0.472	0.469	0.266	0.295	0.081	0.081	0.071	0.080
2016	0.466	0.463	0.270	0.284	0.081	0.081	0.075	0.081
2017	0.470	0.467	0.293	0.270	0.082	0.082	0.094	0.083
2018	0.478	0.475	0.312	0.254	0.084	0.084	0.115	0.084
2019	**0.497**	**0.494**	0.328	0.240	**0.088**	**0.087**	0.131	0.084
2020	**0.537**	**0.534**	0.343	0.225	**0.094**	**0.094**	0.148	0.085

Female Condition

The best-fit model did a poor job of predicting recruitment in 2005-2007 and in 2009. A previous analysis of condition of female sablefish noted that female condition was low in these years (Haltuch et al. 2019b). Since evaluating condition requires individual length-weight data, it has some limitations for hindcasting to data-poor years, so we do not evaluate it in the main manuscript. However, incorporating condition may help to elucidate the model failures above, and we examine its effects on model fit here.

Female sablefish mature at approximately seven years (50% mature at 6.86 years; Head et al. 2014). Therefore, we evaluated whether adding condition for age-7+ females improved the model fit for the years 2003-2019—the years for which condition data (length and individual biomass) were available from the West Coast Groundfish Bottom Trawl Survey (WCGBTS, Keller et al. 2017). We used relationships for females north of Cape Mendocino (40° N) because the sea-level index in the best-fit model was northern sea level, because growth rates differ north and south of Cape Mendocino (Head et al. 2014), and because the majority of the length-age data are from the northern portion of the range (generally May – September for data north of Cape Mendocino) (Haltuch et al. 2019b; Kapur et al. 2021). The condition index (CI) is a relative measure of the overall health of the fish quantified as the observed weight of an individual relative to the expected weight from the length-weight relationship for the species (Ricker 1973, Ricker 1975, Stevenson and Woods 2006). We used data from the WCGBTS to calculate the condition index for female sablefish. We calculated condition for age-7+ females. First, we calculated the length-weight relationship as:

$$\log(W_i) = \log(a) + b*\log(L_i)$$

Where W = weight in kg, L = length in cm, *a* and *b* are estimated parameters, and *i* indicates the individual fish. There was a strong relationship on the log-scale (r^2 = 0.98, Fig. S5).

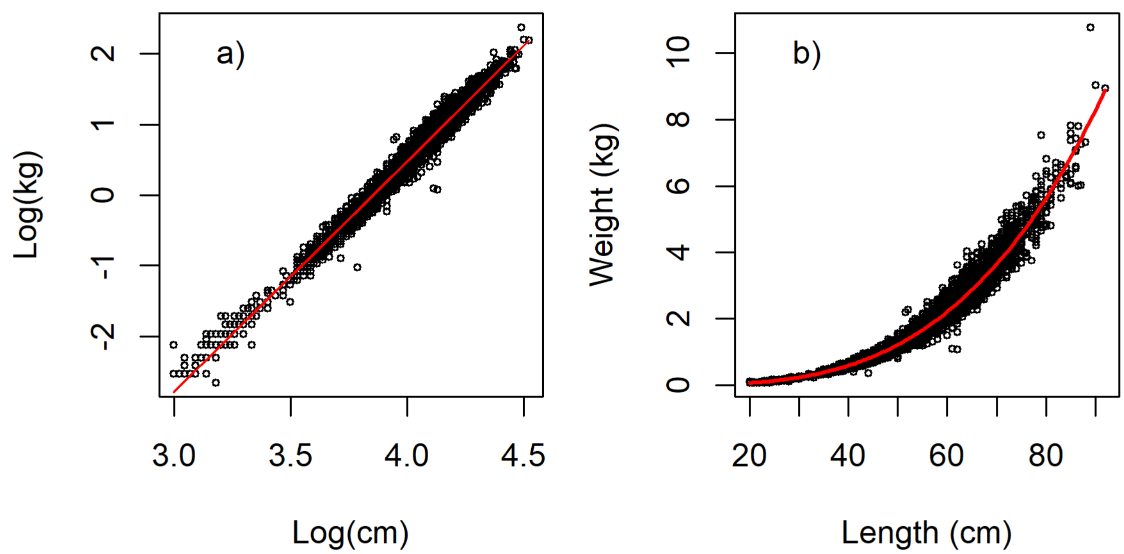

Fig. S8. Length-weight relationships for female sablefish, coast-wide. a) log-scale relationships and b) untransformed data.

Next, we back-transformed the resulting relationship (equation) to the original data scale to obtain the length-weight relationship as $W = aL^b$, where a = 3.30 x 10^{-6}, and b = 3.27. We then calculated condition for each individual as:

$$CI = W_{obsserved}/W_{expected} * 100$$

Finally, we averaged the Individual Condition Index by year to obtain an annual index of female condition for age 7+ females north of approximately Cape Mendocino.

We added female condition to the base model (DF1) in several forms and selected the best-fit model based on the lowest AICc. We add female condition as a continuous variable and as a categorical predictor in which years with condition exceeding the upper and lower 1.0 s.d. bound were classified as “good” or “poor” and other years were classified as “normal” (Fig. S6). We also fit each as lagged one year or estimated recruitment and condition in the same year.

Three models had AICc less than 2.0 (Table S4), including the base DF1 only model (r^2 0.27, ranked third). Including current year condition as a continuous variable produced the lowest AICc and r^2 = 0.37 (Table S5). Condition as a factor produced the highest r^2 (0.48), and closer examination of the model parameters (Table S6) suggests that years with low condition were important to the model fish (coefficient was different from zero).

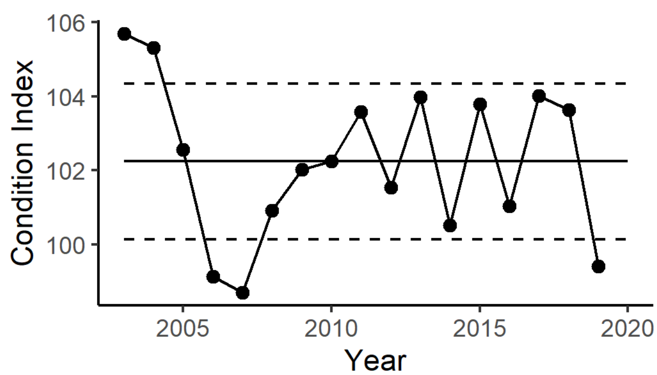

Fig. S9. Condition of age-7+ females north of Cape Mendocino for 2003-2019. See Haltuch et al. (2019b) for details on calculation. Index is the percentage of expected weight for that year. Values below 100 indicate poor condition. Solid line is the average condition across the time series. Dotted lines are +/- 1.0 s.d.

Table S4. Comparison of model fits evaluating predictors of recruitment for 2003-2018 for the base model plus condition as a factor, continuous variable, and lagged or within the same year.

Model	AICc	ΔAICc	R^2	Parameters
DF1 + condition	60.58	-	0.37	3
DF1 + condition, factor	61.30	0.72	0.49	4
DF1	61.48	0.90	0.27	2
DF1 + condition, lagged	62.75	2.17	0.26	3
Condition, lagged	64.04	3.46	0.02	2
Condition	64.37	3.79	0.04	2
DF1 + condition, lagged, factor	64.67	4.09	0.35	4
Condition, factor	64.94	4.36	0.19	3
Condition, factor, lagged	65.30	4.72	0.14	3

Table S5. Results of adding condition (continuous variable) of age 7+ females north of 40° N to the base model predicting recruitment deviations from the stock assessment model.

Parameter	Estimate (SE)	t value	P-value
Intercept	-20.87 (14.11)	-1.479	0.161
Northern sea level (DF1)	-0.93 (0.24)	-2.735	0.016
Condition	0.21 (0.14)	1.491	0.158

Table S6. Results of adding condition (factor = good, average, poor) of age 7+ females north of 40° N to the base model predicting recruitment deviations from the stock assessment model.

Parameter	Estimate (SE)	t value	P-value
Intercept	0.4867 (0.310)	1.572	- 0.1400
Northern sea level (DF1)	-8.8761 (0.319)	-2.744	0.0167
Condition - good	0.3418 (0.829)	-0.412	0.6867
Condition - poor	-1.6030 (0.690)	-2.323	0.0370

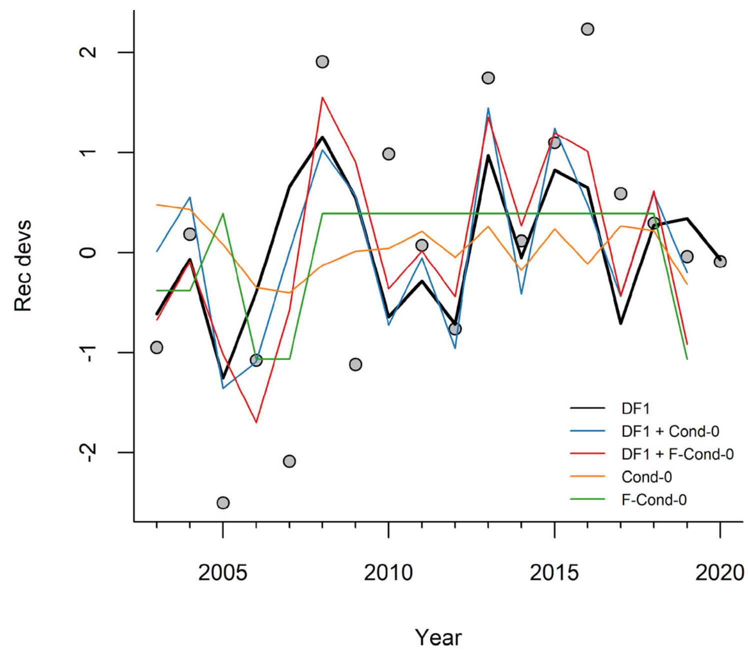

Fig. S10. Results of model fitting using female condition as a categorical variable. 'DF1 Index' are the predicted recruitment deviations for the best-fit model in the main analysis. 'DF1 & F_Cond' are the results when condition is included as a two-level factor in the model. "C_Cond" shows standardized female condition for age-7+ individuals (north of Cape Mendocino), and DF1 is the first dynamic factor from the primary analysis. Dotted lines are the 95% confidence limits for DF1 & F_Cond index.