Received: 6 November 2023 | Revised: 22 May 2024 | Accepted: 2 July 2024

DOI: 10.1111/faf.12850

ORIGINAL ARTICLE

FISH and FISHERIES WILEY

Leveraging ecological indicators to improve short term forecasts of fish recruitment

Eric J. Ward[1] | Mary E. Hunsicker[2] | Kristin N. Marshall[3] | Kiva L. Oken[3] | Brice X. Semmens[4] | John C. Field[5] | Melissa A. Haltuch[3,6] | Kelli F. Johnson[3] | Ian G. Taylor[3] | Andrew R. Thompson[7] | Nick Tolimieri[1]

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

[2]Fish Ecology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Newport, Oregon, USA

[3]Fishery Resource Analysis and Monitoring Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, Washington, USA

[4]Scripps Institution of Oceanography, UC san Diego, La Jolla, California, USA

[5]Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Santa Cruz, California, USA

[6]Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, Washington, USA

[7]Fisheries Resource Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, La Jolla, California, USA

Correspondence
Eric J. Ward, Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Blvd E, Seattle, WA, USA.
Email: eric.ward@noaa.gov

Abstract

Forecasting the recruitment of fish populations with skill has been a challenge in fisheries for over a century. Previous large-scale meta-analyses have suggested linkages between environmental or ecosystem drivers and recruitment; however, applying this information in a management setting remains underutilized. Here, we use a well-studied database of groundfish assessments from the West Coast of the USA to ask whether environmental variables or ecosystem indicators derived from long-term monitoring datasets offer an improvement in our ability to skilfully forecast fish recruitment. A secondary question is which types of modelling approaches (ranging from linear models to non-parametric methods) yield the best forecast skill. Third, we examine whether simultaneous forecasting of multiple species offers an advantage over generating species-specific forecasts. We find that for approximately one third of the 29 assessed stocks, ecosystem indicators from juvenile surveys yields the highest out of sample predictive skill compared to other covariates (including environmental variables from Regional Ocean Modeling System output) or null models. Across

Mary E. Hunsicker, Kristin N. Marshall, Kiva L. Oken and Brice X. Semmens contributed equally to this work.

John C. Field, Melissa A. Haltuch, Kelli F. Johnson, Ian G. Taylor, Andrew R. Thompson and Nick Tolimieri also contributed equally to this work.

modelling approaches, our results suggest that simpler linear modelling approaches do as well or better than more complicated approaches (reducing out of sample Root Mean Square Error by ~40% compared to null models), and that there appears to be little benefit to performing multispecies forecasts instead of single-species forecasts. Our results provide a general framework for generating recruitment forecasts in other species and ecosystems, as well as a benchmark for future analyses to evaluate skill. The most promising applications are likely for species that are short lived, have relatively high recruitment variability, and moderate amounts of age or length data. Forecasts using our approach may be useful in identifying covariates or mechanisms to include in operational assessments but also provide qualitative advice to managers implementing ecosystem based fisheries management.

KEYWORDS
CalCOFI, environmental variables, forecasting, recruitment, ROMS, stock assessment

1 | INTRODUCTION

Interest in forecasting has rapidly increased in fisheries and related fields over the last decade (Dietze et al., 2018), with both data (Hampton et al., 2013) and methodology (Daugaard et al., 2022) evolving to meet complex challenges of making predictions in a non-stationary world (Pennekamp et al., 2017). Examples from aquatic ecosystems include projecting changes in population size (Hare et al., 2010), shifting species distributions (Araújo & New, 2007), biodiversity loss associated with human impacts (Güneralp & Seto, 2013) and disease outbreaks (Woolhouse, 2011). Forecasting quantities of interest relevant to fisheries scientists or managers is inherently a more difficult problem than forecasting in many other fields—data are generally sparse in space and time, the total number of observations may be relatively small, the mechanisms responsible for ecological processes may be incredibly complex and observations are almost always corrupted by measurement and sampling errors (Munch et al., 2018). As a result of these and other factors, complicated forecasting models may do no better in the short term than assuming the future mirrors past conditions (Ward et al., 2014). The challenge then is to identify forecasting approaches, inclusive of analytical method and covariate selection, that improve the prediction of future states even in the face of ecological obfuscation.

There are numerous benefits to being able to generate accurate and precise forecasts in fisheries. These may include improved risk assessments for species of conservation concern (Mace et al., 2008) or better monitoring and management of harvested species (Ovando et al., 2022). Better forecasts may also improve understanding of mechanistic relationships, non-stationary processes, as well as how species or ecosystems will respond to future perturbations. From a methodological standpoint, improved forecasts may also help identify models that are most useful for forecasting, allowing methods to be applied to other systems.

While model-based forecasting approaches have been developed around the world for a wide range of terrestrial and aquatic species, the development of forecasting approaches for marine fishes represents a high risk, high reward venture. Forecasting variable processes, such as fish recruitment (the survival process by which young fish age and become available to a fishery) can be unreliable without a mechanistic understanding of drivers (Hilborn & Walters, 1992). The benefits of including environmental information or other drivers in population models are expected to vary by species and may be marginal for many (Deangelis & Cushman, 1990; Haltuch et al., 2019). Historically, these approaches have been met with a great deal of skepticism when linked to fisheries management (Myers, 1998; Walters & Collie, 1988). That said, the status quo approaches of using forecasts based on estimates of spawning output and the shape of the spawner–recruit relationships or projecting recruitment based on recent conditions are also imperfect, as these relationships have been shown to explain only a small fraction of recruitment variability for most marine fish populations (typically 5%–15% depending on life-history strategy; Cury et al., 2014, Szuwalski et al., 2015). Consequently, linking population recruitment with environmental effects and the ability to provide short-term forecasts were two of five outstanding areas defined by Subbey et al. (2014) as high priority areas for improvement in fisheries assessments. Routine consideration of ecosystem, environmental and socioeconomic drivers of population dynamics is also one of three key priorities for the US National Oceanic and Atmospheric Administration (NOAA) next-generation stock assessment enterprise (Lynch et al., 2018) and is an important priority for other national stock assessment programs as well (Pepin et al., 2020). The scale of fisheries around the world, with an industry valued at more than 400 billion USD (FAO, 2020), offers the ability for even moderate improvements in forecasts to increase ecosystem services and improve social welfare.

There are many examples of quantities in fisheries for which forecasts have been developed—these include recruitment, fisheries catches or landings, population size and distribution and quantities relevant for management (e.g. probability of being below reference points). Using recruitment as a response involves a shorter time window and more opportunity for including external information in forecasts than other forecasted quantities (Haltuch et al., 2019). A challenge in treating recruitment as a response is that it is not directly observable on time scales useful for short-term management decisions but can be estimated within a larger integrated population model (IPM). IPMs combine multiple types of observational data with the aim of improving the precision about quantities of interest, including vital rates, growth rates or population abundance. The use of IPMs is more recent in ecology (Schaub & Abadi, 2011) but these models have been used in fisheries for 50 years (DeFilippo et al., 2021; Maunder & Punt, 2013). Using IPM model output as a response includes unique challenges and requires familiarity with the assumptions in the statistical models used (Brooks & Deroba, 2015). Within the context of fisheries applications, IPMs generally assume that recruitment is stochastic and is governed by a specified stock–recruit relationship (Beverton Holt, Ricker, etc.; Hilborn & Walters, 1992, Maunder & Punt, 2013). Recruitment deviations represent anomalies from the stock–recruit relationship and are independent of stock abundance or biomass. Previous work linking the environment with population dynamics has largely focused on using environmental indicators as predictors of recruitment or recruitment deviations indices (Gross et al., 2022; Haltuch & Punt, 2011; Tolimieri & Haltuch, 2023).

Despite the decades of research attempting to improve forecasts of fish recruitment, several key questions remain. A first question is whether ecosystem indicators offer any advantages over information on the physical environment (e.g. temperature, oxygen and salinity) in improving forecasts of fish recruitment. Indicators may be derived from a number of sources, including abundance of prey, predators or similar species. One of the more promising types of data to improve forecasts of fish recruitment are long-term data collected to monitor densities of larval or pelagic juvenile fishes (Koslow & Wright, 2016; Litzow et al., 2022; Southward et al., 2004; Walsh et al., 2015). A number of previous efforts have linked larval or juvenile fish densities to climate variables (Ibaibarriaga et al., 2007; Lynam et al., 2004; Marshall et al., 2019; Schroeder et al., 2019), and, in some cases, indices of larval or juvenile densities from fisheries surveys have been included in fisheries stock assessment models as indices of spawning output (Adams et al., 2019; Dick & He, 2019; He & Field, 2018). However, these previous approaches have not evaluated whether other species may provide better indices of recruitment or whether multiple indicators may be used to predict recruitment. Species may be useful proxies for one another if, for example, they have similar life-history characteristics or similar responses to environmental variability (Doyle et al., 2009).

A second key question is whether forecasts of fish recruitment may be improved by conducting simultaneous predictions of recruitment for multiple species at the same time. Large-scale synchrony across species has been observed in some systems (Stachura et al., 2014; Zimmermann et al., 2019), and in laboratory settings, forecast skill has been shown to improve with including interspecific interactions (Daugaard et al., 2022). Previous research on wild populations has demonstrated that forecasts may be improved by using data from other species in the same area (Minto et al., 2014; Thorson et al., 2013); however, data from other species may act as red herrings (multispecies interactions may be confounded with changes in the environment), and it remains unclear whether multispecies forecasts may outperform single species ones.

The first objective of this paper is to test whether the inclusion of ecosystem information (larval or pelagic juvenile fish surveys and physical ocean variables) improves forecast skill of future recruitment of fish populations. As some species may respond similarly to environmental perturbations, a second objective is to evaluate whether forecasting recruitment deviations for multiple assessed species at the same time offers advantages compared to single species approaches. For both objectives, ancillary goals are to identify which types of forecasting model(s) perform best and whether particular covariates are consistently useful at forecasting recruitment for multiple species. We develop and test recruitment forecasts using a case study of many species in a single region and provide guidance on how future research may operationalize our results in many regions to embed them in population models that are used for management.

2 | METHODS

2.1 | Recruitment time series

We assembled a database of fisheries stock assessments and ecosystem indicators from the California Current Large Marine Ecosystem (USA West Coast); this region was chosen because many archived assessment models ($n = 29$) and data are publicly available and have been the focus of previous forecasting efforts (Bell et al., 2023). The assessments include 21 species (Table 1), as multiple assessments are done for some species based on biogeographic, genetic or management areas (Table 1). Each assessment includes unique data and model assumptions but the estimation framework is the same, described as an integrated, age-structured population model (Maunder & Punt, 2013; Methot & Wetzel, 2013). All assessments also modelled recruitment deviations similarly, using a stochastic Beverton-Holt stock–recruitment function where annual deviations are modelled with a sum to zero constraint and a fixed likelihood penalty (σ_R) that behaves similarly to the standard deviation of a random effect (Methot & Taylor, 2011; Methot & Wetzel, 2013). Because σ_R values differ by species or assessment (Table 1), we standardized all recruitment deviation time series to have a mean of 0 and standard deviation of 1 prior to further modelling (Figure 1). This standardization also allows results to be more easily compared across assessed species.

TABLE 1 Species and stock assessments from the West Coast of the USA used as response variables in our models predicting recruitment variability. Start and end year of the main recruitment deviations for each species shown with the recruitment standard deviation.

Common	Species	Region	Start	End	σ_R
Black rockfish	*Sebastes melanops*	Washington	1950	2011	0.5
Black rockfish	*Sebastes melanops*	California	1957	2011	0.5
Blue and deacon rockfish	*Sebastes* sp. (*mystinus/diaconus*)	California, Oregon	1950	2016	0.5
Bocaccio	*Sebastes paucispinis*	California	1954	2016	1.0
Cabezon	*Scorpaenichthys marmoratus*	Oregon	1980	2015	0.5
Cabezon	*Scorpaenichthys marmoratus*	California north of Pt. Conception	1962	2017	0.5
Cabezon	*Scorpaenichthys marmoratus*	California south of Pt. Conception	1970	2017	0.7
California scorpionfish	*Scorpaena guttata*	California	1966	2015	0.6
Chilipepper rockfish	*Sebastes goodei*	California	1965	2014	1.0
Gopher and black and yellow rockfish	*Sebastes* sp. (*carnatus/chrysomelas*)	California	1978	2018	0.5
Kelp greenling	*Hexagrammos decagrammus*	Oregon	1980	2013	0.6
Longspine thornyhead	*Sebastolobus altivelis*	Coastwide	1944	2012	0.6
Pacific ocean perch	*Sebastes alutus*	Coastwide	1900	2015	0.7
Rougheye and blackspotted rockfish	*Sebastes* sp. (*aleutianus/melanostictus*)	Coastwide	1900	2011	0.4
Shortspine thornyhead	*Sebastolobus alascanus*	Coastwide	1850	2012	0.5
Widow rockfish	*Sebastes entomelas*	Coastwide	1900	2018	0.6
Yelloweye rockfish	*Sebastes ruberrimus*	Coastwide	1889	2015	0.5
Yellowtail rockfish	*Sebastes flavidus*	North of Cape Mendocino	1932	2015	0.4997
Yellowtail rockfish	*Sebastes flavidus*	South of Cape Mendocino	1945	2016	0.77
Canary rockfish	*Sebastes pinniger*	Coastwide	1933	2015	0.5
Darkblotched rockfish	*Sebastes crameri*	Coastwide	1870	2013	0.75
Dover sole	*Microstomus pacificus*	Coastwide	2011	2019	0.35
Lingcod	*Ophiodon elongatus*	North of Cape Mendocino	1889	2020	0.6
Lingcod	*Ophiodon elongatus*	South of Cape Mendocino	1889	2020	0.6
Sablefish	*Anoplopoma fimbria*	Coastwide	1890	2020	1.4
Vermillion rockfish	*Sebastes miniatus*	California north of Pt. Conception	1970	2020	0.6
Vermillion rockfish	*Sebastes miniatus*	California south of Pt. Conception	1965	2020	0.6
Vermillion rockfish	*Sebastes miniatus*	Oregon	1961	2020	0.6
Vermillion rockfish	*Sebastes miniatus*	Washington	1949	2020	0.6

2.2 | Larval and juvenile fish time series

We used two long-term larval and pelagic young-of-the-year (YOY) fish datasets to generate potential indicators of ecosystem processes (as predictors of groundfish recruitment). To account for spatial variability in sampling from each survey, we applied index standardization methods to generate time series of densities for each sampled species. The California Cooperative Oceanic Fisheries Investigations (CalCOFI) survey has been conducting regular oceanographic and biological sampling, which includes sampling of ichthyoplankton (larval fishes) in the southern range of the California Current since 1951. Originally developed to understand the oceanographic and ecosystem processes that contributed to the decline of the California sardine (*Sardinops sagax*) population, this long-term monitoring program has been widely used to characterize ecosystem processes in the region (Koslow et al., 2013; Koslow & Wright, 2016). We restricted CalCOFI data to recent years with consistent stations (1985–2022) and further restricted filtered species to only include species with observations in 30 or more years and 300 or more total samples (Figure S1 and Table S1). We applied index standardization methods in R (R Core Development Team, 2022) using a spatial generalized linear mixed model with the package sdmTMB (Anderson et al., 2022); additional details on this approach are included in the Supplementary Information.

The pelagic YOY fish data was from the Rockfish Recruitment and Ecosystem Assessment Survey (RREAS) that has been conducted in Central California since 1983 and throughout California waters since 2004 (Field et al., 2021, Santora et al., 2021; Figure S1). This survey uses a pelagic midwater trawl to sample pelagic juvenile stages of groundfish and other forage species (such as krill and market squid), as well as collects oceanographic and predator (seabird and marine mammal) data. Data were subset to include only stations within the core area in Central California and well sampled species or those that have been previously used as indicators (Table S2). Similarly to the CalCOFI data, we generated indices of density separately for each species; additional details are included in the Supplementary Information. Indices were constructed both for the time period where data are publicly available (1990–present), and as a sensitivity analysis, secondary indices were constructed using the entire dataset (1983–present).

2.3 | Environmental and ecosystem data

Following previous work in the California Current (Hunsicker et al., 2022), we used physical oceanographic time series derived from a Regional Ocean Modeling System (ROMS) model for the region (Neveu et al., 2016). Using ROMS output, monthly time series were generated for the following six variables: sea surface temperature (SST), sea surface height (SSH), isothermal layer depth (ILD), Brunt-Väisälä frequency (BV), a coastal upwelling transport index (CUTI) and a biologically effective upwelling transport index (BEUTI)

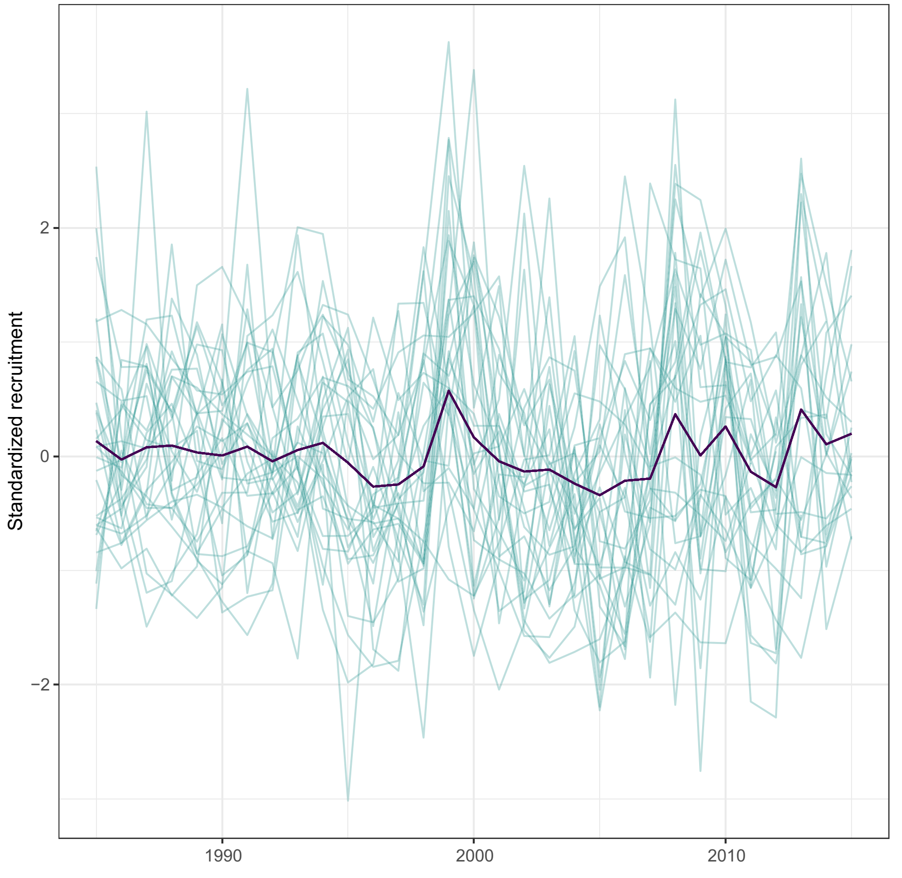

FIGURE 1 Standardized recruitment and global mean (dark line) for the 29 time series of stock assessment recruitment deviations in our analysis.

(Jacox et al., 2018). Following Hunsicker et al. (2022), monthly data were averaged annually (July–June) and calculated separately for a central and southern portion of the California Current separated at Point Conception, California (34.5° N).

In addition to the larval and pelagic juvenile fish time series derived from the CalCOFI and RREAS datasets, respectively, or the environmental time series derived from the ROMS model, we also wished to evaluate whether ecosystem state indices outperform individual time series. We adopted a dimension reduction approach known as Dynamic Factor Analysis (Hunsicker et al., 2022; Zuur et al., 2003) and applied DFA separately to each of the three data sources (CalCOFI, RREAS and ROMS) to generate latent ecosystem trends. Because of the differing number of time series from each dataset, and potentially different numbers of processes generating data for each, we did not combine all time series into a single DFA model. DFA trends from these analyses were then treated as predictor variables (Figures S1 and S2). Further details can be found in the Supplementary Information.

Given that there is only modest overlap between species for which stock assessments are conducted (Table 1) and those for which pelagic juvenile density estimates exist (Tables S1 and S2; Field et al., 2021), the pelagic juvenile fish data can be generally viewed as representing latent ecosystem processes. Specifically, variation in pelagic juvenile fish densities, or that of other pelagic forage species, may track important oceanographic processes, including those that may take place over finer spatial or temporal scales, that are not captured by other covariates, such as the ROMS variables. Comparing models that include ROMS to those that include juvenile fish density allows us to evaluate whether ecological indicators may provide more utility in forecasting recruitment of commercially important fish populations. By examining time series of juvenile fish densities across a large number of species, we can ask whether there are consistent species that may be useful as indicators (across species being forecast and across types of forecasting models used). Finally, by comparing a range of statistical models used for forecasting recruitment, we can identify approaches that may have utility in being included in larger integrated population models.

2.4 | Forecasting recruitment

As a null model, we first generated 1 year forecasts for each of the most recent 10 years of data (data collected through 2005 were used to forecast recruitment in 2006, data through 2006 were used to forecast recruitment in 2007, etc., so that forecasts were generated for holdout years 2006–2015). Our null model consisted of using the historical mean preceding each holdout point (equivalent to fitting an intercept only model). As standardization was done only once for the entire dataset (e.g. 1985–2015), the intercept is expected to be close to but not exactly zero.

We next constructed univariate statistical models to forecast recruitment 1 year into the future. As the temporal overlap between predictors and response time series is relatively short (~30 years in the best cases), we restricted the range of models considered to include the following: linear regression with one to three predictors, generalized additive models (GAMs implemented in 'mgcv'; Wood, 2003) with one to three predictors (using P-spline smooths; Eilers & Marx, 1996), regularized regression (implemented in the 'glmnet' package; Simon et al., 2011) and random forests (implemented in the 'randomForest' package; Liaw & Wiener, 2002). Regression models and GAMs can be viewed as linear mixed-effects models that incorporate smooth effects of covariates. Regularized regression is a type of penalized regression (lasso regression being a special case, where the penalty is proportional to the absolute value of the regression coefficient, and many resulting coefficients are zero). Finally, random forests were included as a machine learning approach, to allow for a potentially larger number of predictor variables, as well as the interactions between them.

As both the predictors and responses are standardized, our parametric models omitted intercepts and only estimated slopes or smooths for each covariate. We conducted a grid search across tuning parameters for regularized regression (elastic-net penalty $\alpha=.1$–1.0, $\lambda=0$–2) and random forest models (number of trees = 300–2000, number of variables selected at each split = 2–10). Though our index standardization of larval and juvenile fish time series generated standard errors, this uncertainty was not included in forecasts because of complications in including it in non-parametric models.

Our univariate statistical models were applied to each combination of recruitment deviations and covariates. For the parametric models (e.g. time series from one to three species from the CalCOFI survey used to forecast recruitment for kelp greenling), we constructed forecast models for each combination of up to three potential covariates (as some drivers are correlated, we summarized the maximum absolute correlation among predictors as a measure of collinearity; Supplementary Information II). For the regularized regression and random forest models, all potential covariates were used simultaneously.

To evaluate support for improved forecasting of recruitment time series with multiple species (Daugaard et al., 2022), we created multivariate extensions of each of the univariate models described above where recruitment deviations of all of the species were modelled simultaneously. A large difference between multivariate models and univariate ones is that when combined in a single model, the assessment can be included as a covariate (factor or random effect) to allow trends to vary across assessments. In addition to the predictors used in the univariate models (e.g. 1–3 time series of covariates for regression models), our parametric multivariate models included the interaction between assessments and covariates either as fixed factors or random effects. The differences between these approaches are subtle, when interactions are treated as factors, no common distribution of effects is assumed, whereas when interactions are random, a common distribution is estimated. Multivariate linear mixed-effects models were implemented using the R package 'glmmTMB' (Brooks et al., 2017). For multivariate GAMs, we modelled covariate effects as hierarchical factor smooths (Pedersen et al., 2019), where the effect of a covariate is allowed to vary uniquely across each assessment (but these relationships have the same wiggliness, modelled again with P-splines; Eilers & Marx, 1996). Our multivariate extension of regularized regression and random forest models included the interactions between assessments and covariates as predictors. Additional details and equations may be found in the Supplementary Information. All code to replicate these analyses is included in our GitHub repository https://github.com/ecosystem-state/recruit-devs; code for applying the univariate and multivariate forecast methodology is bundled in a separate publicly available R package for future potential uses (https://github.com/ecosystem-state/predRecruit).

2.5 | Quantifying performance

We used the out-of-sample predictions over the period 2006–2015 to calculate the root mean squared error (RMSE) as a measure of future predictive skill (leaving out all future information). RMSE measures predictive skill in absolute terms, incorporating the bias and variance of predictions. RMSE was calculated separately for each combination of responses, covariates and modelling approaches used. In addition to RMSE, we calculated R^2 over the forecast period as a measure of the proportion of variance explained. As a more qualitative measure of predictive skill, we discretized the time series of recruitment deviations and predictions into terciles ('low', 'average' and 'high' recruitment) using quantiles from a standard normal distribution to define thresholds. Following Kiaer et al. (2021), we then calculated the hit rate (proportion of correct tercile forecasts) and true skill score (hit rate minus false alarm rate). As the error rates may not be symmetric, and fishery managers may care more about predicting and avoiding worst-case scenarios, we also examined the forecast skill of correctly predicting each tercile bin ('low', 'average' and 'high'). Finally, to calculate the degradation in forecast skill with increasing forecast horizons, we compared the RMSE of 1- and 2-year forecasts to a baseline of RMSE applied to the training data alone.

2.6 | Calculating variable importance

Many types of variable importance scores can be computed for the classes of models considered in our analysis (Grömping, 2009). Focusing on the parametric models (linear models, GAMs), we calculated variable importance scores as the mean marginal improvement in RMSE. Marginal improvements in RMSE are dependent on both information in a predictor, as well as redundancy in information across predictors (e.g. if variable x and y are in a model, and z is highly correlated with x, adding z to a model will not improve RMSE). For any potential covariate considered in models with m predictors, we used the set of n models with $m-1$ predictors that did not include that covariate as a baseline, and for each calculated the average relative change in RMSE

$$p_i = \frac{1}{n}\sum_{j=1}^{n}\frac{\mathrm{RMSE}_{j,m-1}-\mathrm{RMSE}_{j,m}}{\mathrm{RMSE}_{j,m-1}}$$

where $\mathrm{RMSE}_{j,m-1}$ might represent RMSE from a model with 2 predictors, and $\mathrm{RMSE}_{j,m}$ the RMSE from a model containing the original covariates plus a third predictor added. As we were interested in also summarizing the marginal improvement across varying numbers of predictors (1–3), we then calculated the average $\sum_{i=1}^{3} p_i/3$ for each predictor variable.

3 | RESULTS

3.1 | Which covariates are most useful?

For the majority of stocks included in our analysis, the addition of covariate data (ROMS, CalCOFI, RREAS and DFA trends) improved the predictability over a null model with no covariates (Figures 2 and 5). Comparing the relative performance of individual covariate datasets, we found that larval and pelagic juvenile fish data—especially those generated from the CalCOFI survey—reduced RMSE scores and increased the R^2 of recruitment forecasts (models using CalCOFI drivers performed best in 20 of 29 assessments; Figures 2 and 5 and Figures S4 and S5). Across assessments, the best individual models using CalCOFI drivers resulted in an average 38% decrease in RMSE for linear models and a 41% decrease in RMSE for GAMs, compared to a null model (Figure 2). Indices derived from the RREAS survey also performed well for many species and performed similarly to CalCOFI indices for several stocks (Bocaccio, Blue/Deacon rockfish, Figure 3 and Figure 5). In our sensitivity analysis using RREAS data from 1983 to present versus 1990 to present, we found similar results for the majority of stocks (Figures S9 and S10). Compared to the original data, indices generated via dynamic factor analysis (DFA) generally did not perform as well (Figure 2 and Figure 5).

3.2 | Which modelling approaches are most useful?

For all 29 assessments included in our analysis, parametric and semi-parametric approaches (linear models, GAMs) had lower out-of-sample RMSE values compared to lasso regression or random forests (linear models and GAMs resulting in the best RMSE in 29 of 29 assessments; Figure 2). The clearest cases of linear models and GAMs outperforming other approaches are when larval density estimates from CalCOFI are used as predictors (lingcod, Pacific Ocean perch, sablefish, yellowtail rockfish, widow rockfish; Figure 2 and Figure 5). As expected, we found an inverse relationship between RMSE and R^2 values; approaches that decreased RMSE generally also increased R^2 values (Figures 2 and 5 and Figures S4 and S5). Though our focus is on forecasting recruitment deviations, methods that performed

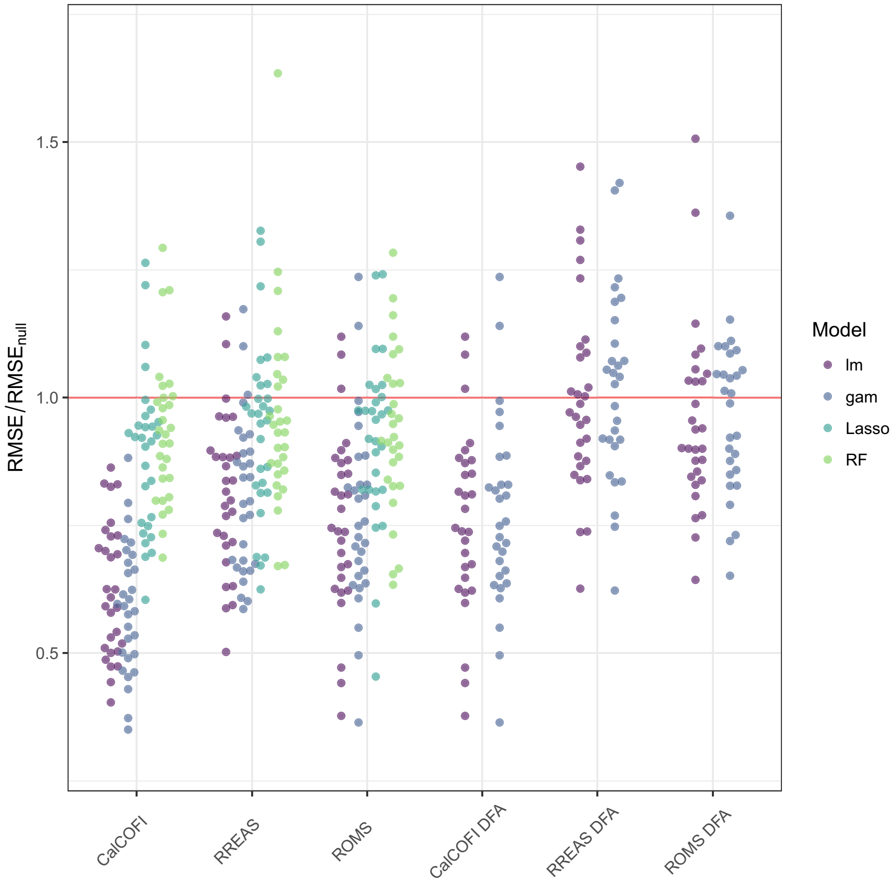

FIGURE 2 Ratio of root mean square error (RMSE) from 1-year forecasts of recruitment deviations of the best model for each combination of predictor variables (*x* axis) to RMSE derived from a null model with no covariate, modelling approaches (colors) and stock assessment time series (points) used as a response. Points below the horizontal red line at 1.0 indicate cases where the forecast outperforms the null model.

poorly for out-of-sample predictions (e.g. random forest) tended to have superior in-sample performance when only evaluated on training data (Figure S6). Interpreting qualitative results in forecasting low or high terciles was difficult (Figure S7). Many of the error rates were at or near 100% because of small sample sizes overall ($n=10$ points with forecasts), as well as the number of values of each tercile over the most recent decade (0–10).

Focusing on results for the linear regression and GAM models, forecast skill for all assessments showed a decrease in performance for 2-year versus 1-year forecasts. However several species showed similar predictive skill 2-year ahead forecasts (lingcod, longspine thornyhead, widow rockfish; Figure S8). Comparing modelling approaches, there was little difference between linear regression and GAMs for 2-year versus 1-year forecasts (bocaccio, canary rockfish, widow rockfish) but for others, linear regression did substantially better for the 2-year forecast horizon (Cabezon, Rougheye/blackspotted rockfish; Figure S8).

3.3 | Which species are most useful in forecasting?

Focusing on eight assessments with high R^2 values (>.75; Figure 3; Supplemental Information II), we found that larval fish densities from CalCOFI offered the largest improvements relative to other covariates (Figure 3). Recruitment deviations from the southern lingcod assessment appear to be highly linked to larval myctophid abundance (e.g. lantern fish) and larval abundance of other forage-fish species, while bocaccio recruitment deviations were correlated with densities of larval sanddabs and viperfish (Figure 3). Scorpionfish and longspine thornyhead recruitment deviations represented two cases where ROMS variables (isothermal layer depth and temperature) outperformed larval and pelagic juvenile fish indices.

3.4 | Do multivariate approaches improve on single species models?

We used a subset of data (recruitment deviations up to and including 2011) to examine whether multivariate extensions of our univariate models offer any improvement in predictive capability. The largest difference between these approaches is the potential to include pooled relationships (e.g. random effects) or assessment–covariate interactions (regularized regression and random forest). For the fixed effect regression and GAM approaches, univariate models resulted in lower RMSE values, higher hit rates and higher total skill scores (hit rate – false rate) when predicting terciles of future recruitment deviations (Figure 4). The most promising results with the multivariate model were for linear models with partial pooling (i.e. GLMM; Figure 4)—for all covariate datasets (CalCOFI, RREAS and ROMS), these models resulted in the lowest RMSE scores. Advantages of the

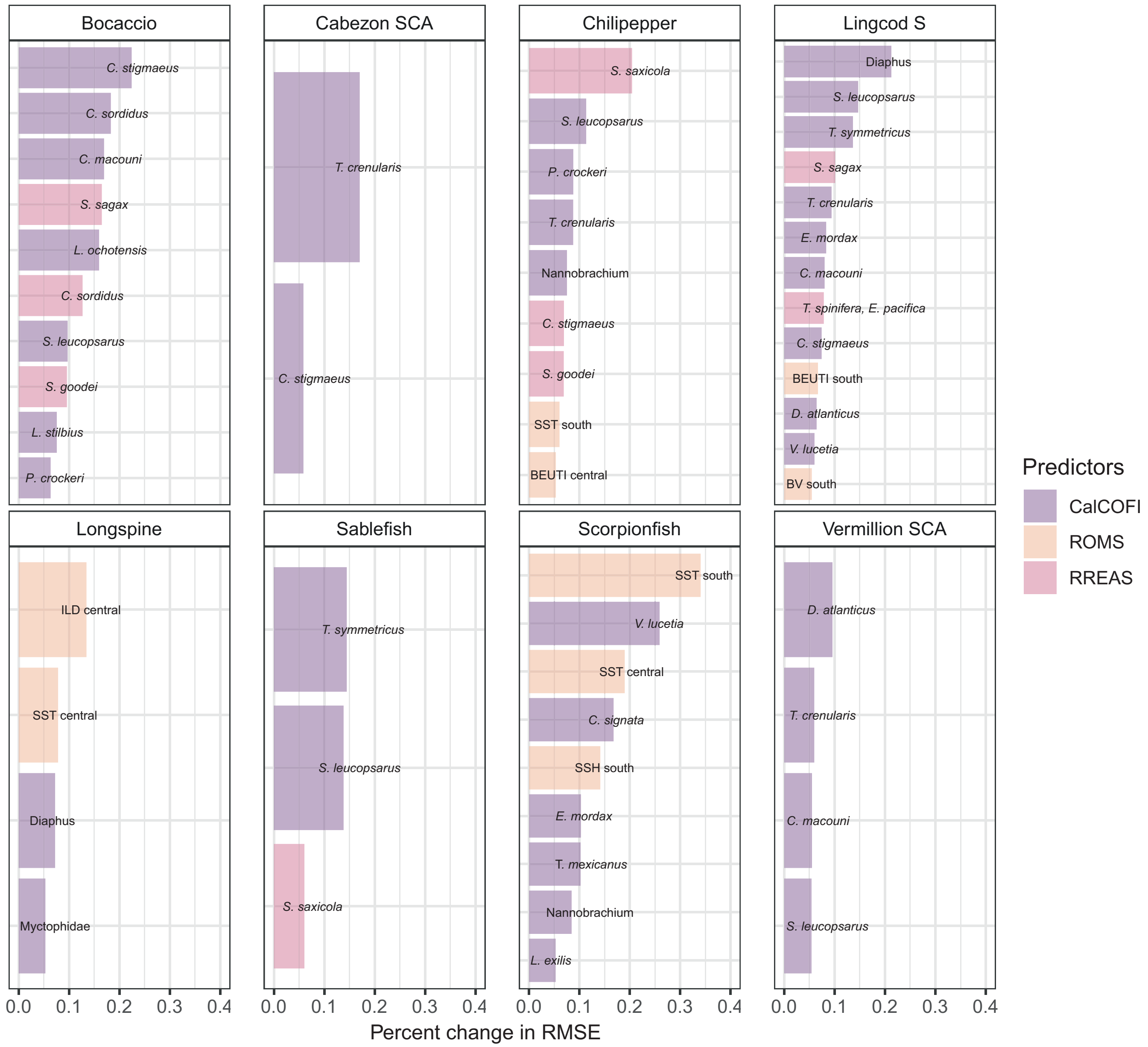

FIGURE 3 Relative variable importance for linear models; each panel represents a stock assessment (Lingcod S is the assessment for the stock south of Cape Mendocino, Cabezon SCA and Vermillion SCA represent the assessments done in southern California; Table 1). The average percent change in root mean square error (RMSE) is calculated as the marginal improvement from two- to three-parameter models (e.g. 0.3 represents a 30% improvement in RMSE).

multivariate approach were less clear for the regularized regression and random forest models, which were associated with the highest RMSE values (Figure 4).

4 | DISCUSSION

Understanding and integrating ecosystem information into fisheries management models is paramount in achieving sustainable fisheries and preserving marine biodiversity. The role of environmental variability in influencing recruitment has been established in ecosystems around the world (Mäntyniemi et al., 2013; Moustahfid et al., 2021; Stige et al., 2013). Environmental covariates have been included in several fisheries stock assessment models (Kapur et al., 2021; Pepin et al., 2020; Stock & Miller, 2021) and also used as predictors in models evaluating long-term change in productivity of marine populations (Britten et al., 2016). In the face of rapidly changing climates, it is imperative that environmental information also be used in generating short-term forecasts, giving fisheries scientists, managers and stakeholders who are grappling with the challenge of managing and protecting marine resources a preview of future conditions on time scales that are relevant for decision making.

To help advance these efforts, we used an established dataset of fisheries stock assessment models and long-term monitoring

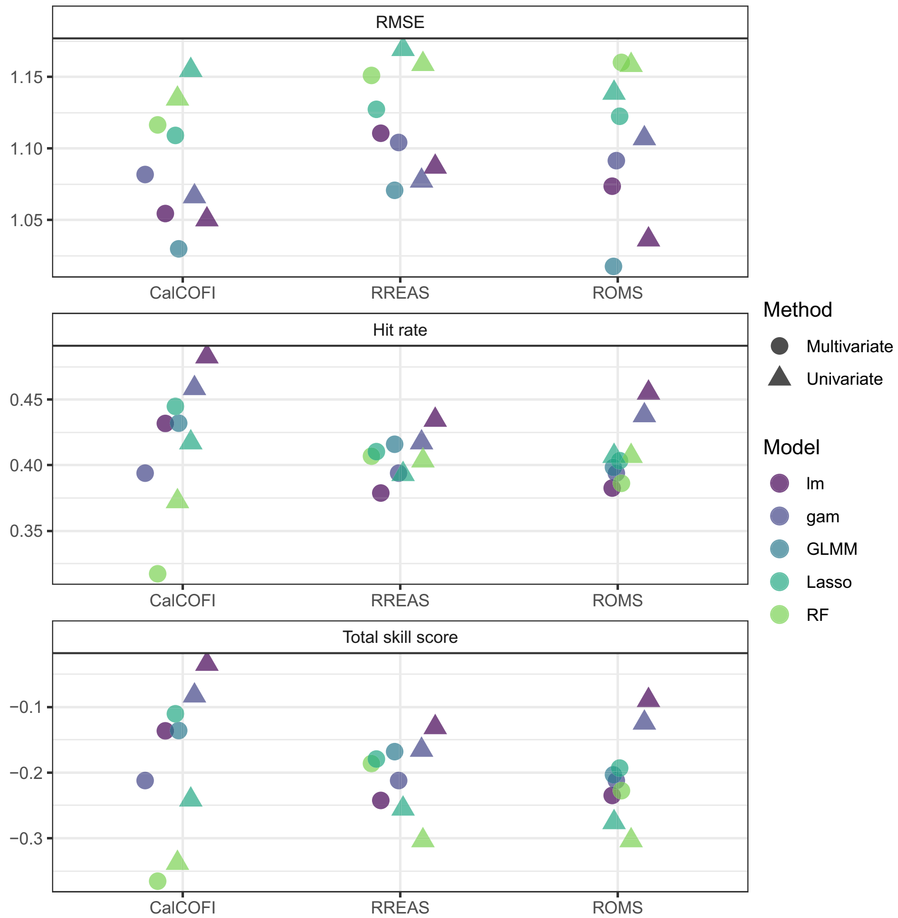

FIGURE 4 Comparison between univariate and multivariate forecasting approaches, using root mean square error (RMSE), the hit rate and total skill score (hit rate − false alarm rate) as performance metrics (*y*-axis). To allow comparison between univariate and multivariate models, results are combined across all assessments, using data through 2011.

datasets from the West oast of the USA to examine the utility of using environmental or ecosystem processes to provide skilful short-term forecasts of recruitment. All of these datasets have previously been used to characterize the ecosystem response to oceanographic variability (e.g. Koslow et al., 2013; Ralston et al., 2015; Schroeder et al., 2019), and our analysis was based on the assumption that the oceanographic or ecological processes and mechanisms responsible for the variability in these datasets are similar or consistent with those that are driving the year-to-year variation in recruitment or year-class strength indicated by the stock assessment models. Though these assessments are produced by a single modelling framework, many of the insights from our analysis can be generalized to other frameworks.

For each of the 29 stock assessments included in our analysis, we evaluated forecast skill across multiple statistical methods (linear regression, GAMs, regularized regression and random forest) and covariates (larval fish densities from the CalCOFI survey and juvenile fish densities from the RREAS survey and ROMS oceanographic variables). Notably, our analyses highlighted that in many cases, the inclusion of external covariates improved the forecast skill (reduced RMSE). Some of the biggest advantages were for recruitment of assessed stocks inhabiting the southern end of the California Current (Figure 3) with larval fish indices from the CalCOFI survey used as predictor variables. This result is intuitive given that CalCOFI sampling is the most robust in this region. Across modelling approaches, simple linear regression approaches appeared to generally outperform more complicated non-linear and non-parametric methods. This result appears largely consistent with previous forecasting competitions, where simpler linear models have outperformed more complicated non-parametric ones (Ward et al., 2014). On a positive note, implementing these models within existing assessment frameworks will introduce fewer hurdles than machine learning techniques.

As expected, we found that the predictive skill of our short-term forecasts degraded from 1 to 2 years for all species (Figure S8; skill decreases for Pacific Ocean perch and cabezon). To robustly evaluate predictive skill, we used multiple metrics to evaluate forecasts (RMSE, hit rates), as each is expected to yield different inferences about the predictive performance of alternative models or covariates. Further, errors may also be asymmetric, where costs associated with wrongly forecasting normal or high recruitment in a bad year are worse than wrongly forecasting poor recruitment in a normal or high year.

4.1 | Challenges of different data across assessments/species

Though not necessarily the cases with the most useful forecasts, the eight assessments with the highest out of sample R^2 (>.75)

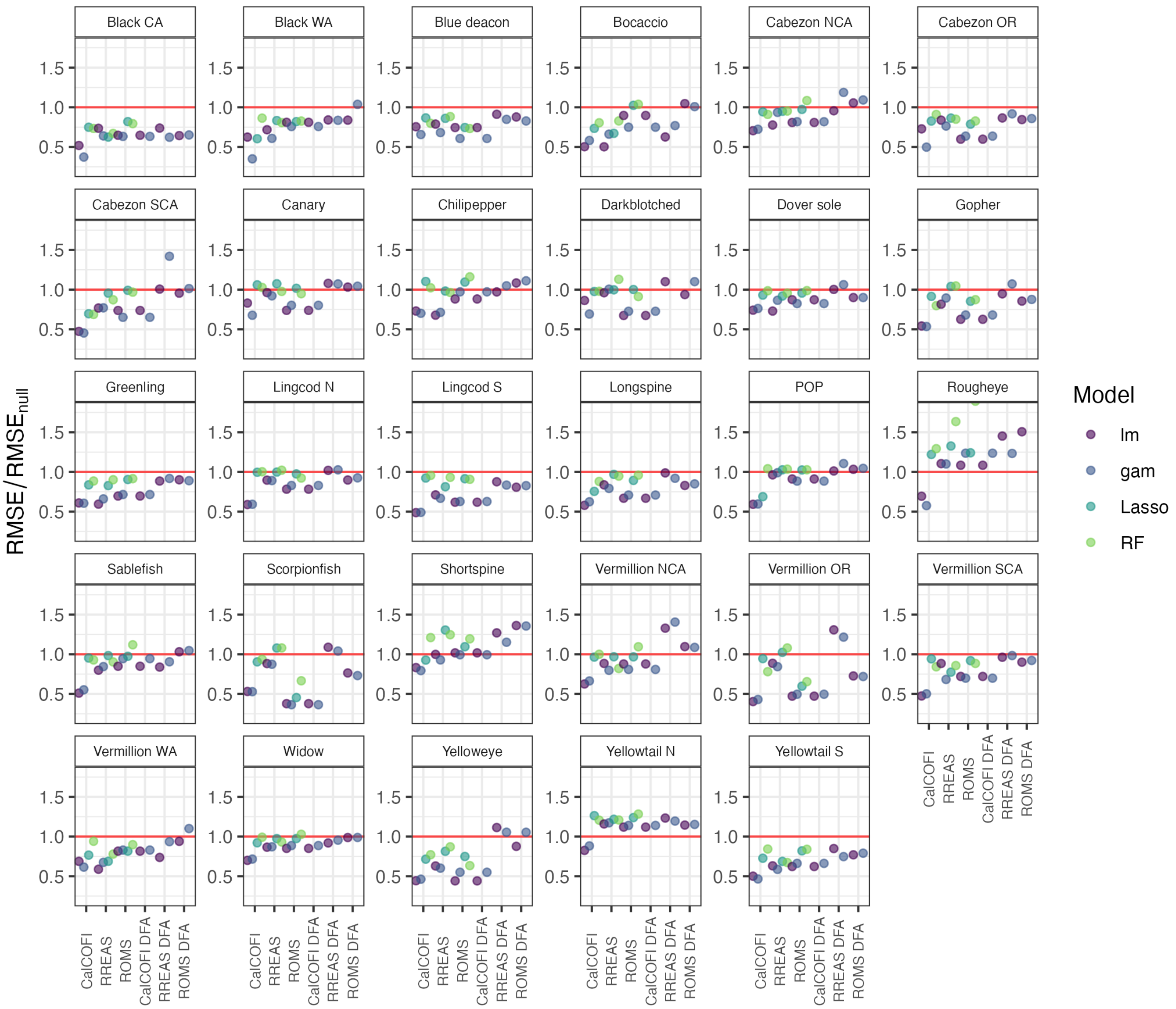

FIGURE 5 Ratio of root mean square error (RMSE) from 1-year forecasts of recruitment deviations of the best model for each combination of predictor variables (*x* axis) to RMSE derived from a null model with no covariate, modelling approaches (colors) and stock assessment time series (points) used as a response. Points below the horizontal red line at 1.0 indicate cases where the forecast outperforms the null model. Each panel represents a stock assessment (Lingcod S is the assessment for the stock south of Cape Mendocino, Cabezon SCA and Vermillion SCA represent the assessments done in southern California; Table 1).

represent a mix of stocks, with a generally southern distribution. One challenge in using our results to improve mechanistic understanding of drivers is that these assessments represent a mix of spatial resolution (Figure 3). At one extreme, stocks like sablefish are distributed from Alaska to California and assessments combine data from multiple regions (e.g. states of Washington, Oregon and California). Being a transboundary stock, ocean conditions in more northern regions can influence sablefish population dynamics throughout the range. At the other extreme, several species in our analysis have separate stock assessments for each region (e.g. cabezon and vermillion rockfish are assessed separately in Oregon and Northern and Southern California). These breaks are generally based on biological information related to estimated or assumed stock structure, with more nearshore stocks typically found or assumed to have more resolved stock structure and dynamics than deeper water species (Gunderson & Vetter, 2006). While our analysis used the same ROMS oceanographic covariates as predictors for all recruitment time series, future work might improve on this by generating stock and area-specific indices based on breaks used in the assessment process, known depth ranges (Haltuch et al., 2020) or genetic structure (Longo et al., 2020). Combining specific species with shared geographic boundaries and/or life-history characteristics may also improve on the utility of multispecies models.

We expect that the inclusion of environmental covariates within assessments will have the most impact on recruitment estimates during the last years modelled within an assessment (where recruitment deviations are generally less precise due to the lack of demographic data to inform the estimates) and for stocks with a lack of data providing information about recruitment deviations (e.g. stocks without larval/juvenile fish surveys or fisheries that target smaller, younger fish; Tolimieri & Haltuch, 2023). Similarly, we expect the inclusion of environmental covariates to differ across species and have a greater impact on those that are shorter-lived and have greater recruitment variability (Moustahfid et al., 2021).

A variety of biological, fishery and management factors impact both the amount and information content of age- and length-composition data in assessments, and thus how reliable estimates of recruitment are. At one extreme in our case study, longspine thornyhead is a long-lived and slow-growing species, with little age data and no validated ageing method (Stephens & Taylor, 2013). At the other extreme, the assessment of sablefish generally includes over 1000 aged individuals per year from fishery-independent surveys alone, and young fish are encountered by both surveys and fisheries, allowing for robust estimation of year class strength within just a few years of recruitment. For models with little or no age data, recruitment (and thus recruitment deviation) estimates rely on age–length relationships and temporal changes in length-frequency distributions to inform recruitment; these estimates can be robust in rapidly growing fish though often lead to imprecise recruitment estimates or autocorrelated 'blurring' of recruitment deviations across years for slower growing species.

4.2 | Links to management

The time series and statistical methods in our analysis have several potential onramps to fisheries management. First, there is the potential for including ecosystem and environmental covariates directly into assessment models (Kapur et al., 2021; Stock & Miller, 2021). We found that ecosystem indicators derived from larval or juvenile fish surveys were more important predictors than ROMS variables for a number of stock assessments (Figure 3), suggesting that there may be a benefit to evaluating alternative predictors of recruitment for species like sablefish. A future challenge in expanding assessment models to include multiple potential drivers is that covariate selection and model comparison ideally should be embedded within the assessment modelling framework (du Pontavice et al., 2022). There has been a general tendency of correlations between the environment and fish recruitment to break down over time (Myers, 1998), thus it is important that added drivers are informed by mechanistic understanding or causal pathways and that relationships continue to be re-evaluated.

Recruitment forecasts may also be used by managers and stakeholders outside of assessment models. Using the West Coast of the USA as an example, a general limitation of existing assessment procedures for groundfish is that at best, assessments are conducted every 2–4 years (but more commonly, at 8–10-year intervals). For our case study, and assessments around the world with similar updating frequencies, recruitment forecasts may provide a more intermediate picture of stock trends in the temporal gaps between assessments.

The utility of forecasts within fisheries management will vary by species and region. Again, using our US West Coast as an example, there are three primary ways these forecasts could be incorporated into management. First, recruitment forecasts may be used as species-specific or Fisheries Management Plan-specific indicators by the Pacific Fisheries Management Council (PFMC) in annual ecosystem status reports (Figures S11 and S12). Second, recruitment forecasts may be used alongside other metrics (fishery importance and biomass trends) to improve the existing stock assessment framework for prioritizing which stocks are to be assessed in upcoming management cycles. Third, recruitment forecasts could also be used to directly inform the harvest specifications process; forecasts could inform catch levels set through the harvest control rules, inform the magnitude of scientific uncertainty assumed for an assessment (e.g. this uncertainty increases with the age of PFMC assessments), or influence the PFMC's risk tolerance for exceeding the overfishing limit. A risk table approach similar to that proposed by Dorn and Zador (2020) may be a useful framework to translate recruitment forecasts to levels of uncertainty or risk.

4.3 | Future work

Our work represents an important step in linking environmental and ecosystem information with operational assessments of commercially important fish stocks. These results demonstrate that skilful short-term recruitment forecasts can be generated for a number of species on the West Coast of the USA, and our approach offers a new use of long-term larval and juvenile fish monitoring data. These results may be improved with future data collection and modelling efforts. In particular, spatial models of juvenile fish size composition (or growth) may help resolve why some species from these surveys act as reliable indicators of groundfish recruitment, while others do not (Swalethorp et al., 2022). While the availability and type of ecosystem data varies across ecosystems, the methods developed here may be adopted for any available drivers. Additional modelling efforts will also help increase the uptake of environmental data into assessments, including large-scale simulation efforts of assessment models with covariates. Finally, there is a need to explore the utility of environmental covariates across assessment models; collaborations extending across geographic boundaries and assessment model frameworks will ensure robustness of modelling approaches.

DATA AVAILABILITY STATEMENT

All data and code to reproduce our analyses are in a collection of publicly available repositories. Response and predictor data are bundled into our 'recdata' package (https://github.com/ecosystem-state/recdata). Code to reproduce our analyses for the CalCOFI and RREAS surveys are available in the 'calcofi-auto' (https://github.com/ecosystem-state/calcofi-auto) and 'rreas-auto' packages (https://github.com/ecosystem-state/rreas-auto). Finally, the 'predRecruit' package (https://github.com/ecosystem-state/predRecruit) includes wrappers for automating the forecasts of fish recruitment for our analyses. Additional RREAS data not available on ERDDAP is available from John Field, john.field@noaa.gov.

ORCID

Eric J. Ward https://orcid.org/0000-0002-4359-0296
Mary E. Hunsicker https://orcid.org/0000-0002-3036-1515
Kelli F. Johnson https://orcid.org/0000-0002-5149-451X
Nick Tolimieri https://orcid.org/0000-0003-0581-9915

REFERENCES

Adams, G. D., Kapur, M. S., McQuaw, K., Thurner, S., Hamel, A. S., & Wetzel, C. R. (2019). *Stock assessment update: Status of widow rockfish (Sebastes entomelas) along the U.S. West Coast in 2019*. Pacific Fishery Management Council.

Anderson, S. C., Ward, E. J., English, P. A., & Barnett, L. A. K. (2022). sdmTMB: An R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. *bioRxiv*. https://doi.org/10.1101/2022.03.24.485545

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, *22*, 42–47.

Bell, R. J., Tableau, A., & Collie, J. S. (2023). Changes in the productivity of US West Coast fish stocks. *Fisheries Research*, *264*, 106712.

Britten, G. L., Dowd, M., & Worm, B. (2016). Changing recruitment capacity in global fish stocks. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 134–139.

Brooks, E. N., & Deroba, J. J. (2015). When "data" are not data: The pitfalls of post hoc analyses that use stock assessment model output. *Canadian Journal of Fisheries and Aquatic Sciences*, *72*, 634–641.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*, 378–400.

Cury, P., Fromentin, J.-M., Figuet, S., & Bonhommeau, S. (2014). Resolving Hjort's dilemma: How is recruitment related to spawning Stock biomass in marine fish? *Oceanography*, *27*, 42–47.

Daugaard, U., Munch, S. B., Inauen, D., Pennekamp, F., & Petchey, O. L. (2022). Forecasting in the face of ecological complexity: Number and strength of species interactions determine forecast skill in ecological communities. *Ecology Letters*, *25*, 1974–1985.

Deangelis, D. L., & Cushman, R. M. (1990). Potential application of models in forecasting the effects of climate changes on fisheries. *Transactions of the American Fisheries Society*, *119*, 224–239.

DeFilippo, L. B., Buehrens, T. W., Scheuerell, M., Kendall, N. W., & Schindler, D. E. (2021). Improving short-term recruitment forecasts for coho salmon using a spatiotemporal integrated population model. *Fisheries Research*, *242*, 106014.

Dick, E. J., & He, X. (2019). *Status of Cowcod (Sebastes levis)*. Pacific Fishery Management Council.

Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., Keitt, T. H., Kenney, M. A., Laney, C. M., Larsen, L. G., Loescher, H. W., Lunch, C. K., Pijanowski, B. C., Randerson, J. T., Read, E. K., Tredennick, A. T., Vargas, R., Weathers, K. C., & White, E. P. (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., & Zador, S. G. (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., Picquelle, S. J., Mier, K. L., Spillane, M. C., & Bond, N. A. (2009). Larval fish abundance and physical forcing in the Gulf of Alaska, 1981–2003. *Progress in Oceanography*, *80*, 163–187.

du Pontavice, H., Miller, T. J., Stock, B. C., Chen, Z., & Saba, V. S. (2022). Ocean model-based covariates improve a marine fish stock assessment when observations are limited. *ICES Journal of Marine Science*, *79*, 1259–1273.

Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, *11*, 89–121.

FAO. (2020). *The state of world fisheries and aquaculture 2020: Sustainability in action*. FAO.

Field, J. C., Miller, R. R., Santora, J. A., Tolimieri, N., Haltuch, M. A., Brodeur, R. D., Auth, T. D., Dick, E. J., Monk, M. H., Sakuma, K. M., & Wells, B. K. (2021). Spatiotemporal patterns of variability in the abundance and distribution of winter-spawned pelagic juvenile rockfish in the California current. *PLoS One*, *16*, e0251638.

Grömping, U. (2009). Variable importance assessment in regression: Linear regression versus random Forest. *The American Statistician*, *63*, 308–319.

Gross, J. M., Sadler, P., & Hoenig, J. M. (2022). Evaluating a possible new paradigm for recruitment dynamics: Predicting poor recruitment for striped bass (*Morone saxatilis*) from an environmental variable. *Fisheries Research*, *252*, 106329.

Gunderson, D., & Vetter, R. (2006). Temperate rocky reef fishes. In J. P. Kritzer & P. F. Sale (Eds.), *Marine metapopulations* (pp. 69–118). Elsevier Academic Press.

Güneralp, B., & Seto, K. C. (2013). Futures of global urban expansion: Uncertainties and implications for biodiversity conservation. *Environmental Research Letters*, *8*, 014025.

Haltuch, M. A., Brooks, E. N., Brodziak, J., Devine, J. A., Johnson, K. F., Klibansky, N., Nash, R. D. M., Payne, M. R., Shertzer, K. W., Subbey, S., & Wells, B. K. (2019). Unraveling the recruitment problem: A review of environmentally-informed forecasting and management strategy evaluation. *Fisheries Research*, *217*, 198–216.

Haltuch, M. A., & Punt, A. E. (2011). The promises and pitfalls of including decadal-scale climate forcing of recruitment in groundfish stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, *68*, 912–926.

Haltuch, M. A., Tolimieri, N., Lee, Q., & Jacox, M. G. (2020). Oceanographic drivers of petrale sole recruitment in the California current ecosystem. *Fisheries Oceanography*, *29*, 122–136.

Hampton, S. E., Strasser, C. A., Tewksbury, J. J., Gram, W. K., Budden, A. E., Batcheller, A. L., Duke, C. S., & Porter, J. H. (2013). Big data and the future of ecology. *Frontiers in Ecology and the Environment*, *11*, 156–162.

Hare, J. A., Alexander, M. A., Fogarty, M. J., Williams, E. H., & Scott, J. D. (2010). Forecasting the dynamics of a coastal fishery species using a coupled climate–population model. *Ecological Applications*, *20*, 452–464.

He, X., & Field, J. (2018). *Stock assessment update: Status of Bocaccio, Sebastes paucispinis, in the conception, Monterey and Eureka INPFC areas for 2017*. Pacific Fishery Management Council.

Hilborn, R., & Walters, C. J. (1992). *Quantitative fisheries stock assessment*. Springer US.

Hunsicker, M. E., Ward, E. J., Litzow, M. A., Anderson, S. C., Harvey, C. J., Field, J. C., Gao, J., Jacox, M. G., Melin, S., Thompson, A. R., & Warzybok, P. (2022). Tracking and forecasting community responses to climate perturbations in the California current ecosystem. *PLOS Climate*, *1*, e0000014.

Ibaibarriaga, L., Irigoien, X., Santos, M., Motos, L., Fives, J. M., Franco, C., Lago De Lanzós, A., Acevedo, S., Bernal, M., Bez, N., Eltink, G., Farinha, A., Hammer, C., Iversen, S. A., Milligan, S. P., & Reid, D. G. (2007). Egg and larval distributions of seven fish species in north-east Atlantic waters. *Fisheries Oceanography*, *16*, 284–293.

Jacox, M. G., Edwards, C. A., Hazen, E. L., & Bograd, S. J. (2018). Coastal upwelling revisited: Ekman, Bakun, and improved upwelling indices for the U.S. West Coast. *Journal of Geophysical Research: Oceans*, *123*, 7332–7350.

Kapur, M. S., Lee, Q., Correa, G. M., Haltuch, M. A., Gertseva, V., & Hamel, O. S. (2021). *Status of sablefish (Anoplopoma fimbria) along the US west coast in 2021*. Pacific Fishery Management Council.

Kiaer, C., Neuenfeldt, S., & Payne, M. R. (2021). A framework for assessing the skill and value of operational recruitment forecasts. *ICES Journal of Marine Science*, *78*(10), 3581–3591. https://doi.org/10.1093/icesjms/fsab202

Koslow, A. J., Goericke, R., & Watson, W. (2013). Fish assemblages in the Southern California current: Relationships with climate, 1951–2008. *Fisheries Oceanography*, *22*, 207–219.

Koslow, J. A., & Wright, M. (2016). Ichthyoplankton sampling design to monitor marine fish populations and communities. *Marine Policy*, *68*, 55–64.

Liaw, A., & Wiener, M. (2002). Classification and regression by random forest. *R News*, *2*(3), 18–22.

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

Longo, G. C., Lam, L., Basnett, B., Samhouri, J., Hamilton, S., Andrews, K., Williams, G., Goetz, G., McClure, M., & Nichols, K. M. (2020). Strong population differentiation in lingcod (*Ophiodon elongatus*) is driven by a small portion of the genome. *Evolutionary Applications*, *13*, 2536–2554.

Lynam, C. P., Hay, S. J., & Brierley, A. S. (2004). Interannual variability in abundance of North Sea jellyfish and links to the North Atlantic oscillation. *Limnology and Oceanography*, *49*, 637–643.

Lynch, P. D., Methot, R. D., & Link, J. S. (2018). *Implementing a next generation stock assessment enterprise: An update to the NOAA fisheries stock assessment improvement plan*. Page 127. U.S. Department of Commerce, NOAA Tech. Memo. NMFS-F/SPO-183.

Mace, G. M., Collar, N. J., Gaston, K. J., Hilton-Taylor, C., Akçakaya, H. R., Leader-Williams, N., Milner-Gulland, E. J., & Stuart, S. N. (2008). Quantification of extinction risk: IUCN's system for classifying threatened species. *Conservation Biology*, *22*, 1424–1442.

Mäntyniemi, S., Uusitalo, L., Peltonen, H., Haapasaari, P., & Kuikka, S. (2013). Integrated, age-structured, length-based stock assessment model with uncertain process variances, structural uncertainty, and environmental covariates: Case of Central Baltic herring. *Canadian Journal of Fisheries and Aquatic Sciences*, *70*, 1317–1326.

Marshall, K. N., Duffy-Anderson, J. T., Ward, E. J., Anderson, S. C., Hunsicker, M. E., & Williams, B. C. (2019). Long-term trends in ichthyoplankton assemblage structure, biodiversity, and synchrony in the Gulf of Alaska and their relationships to climate. *Progress in Oceanography*, *170*, 134–145.

Maunder, M. N., & Punt, A. E. (2013). A review of integrated analysis in fisheries stock assessment. *Fisheries Research*, *142*, 61–74.

Methot, R. D., & Taylor, I. G. (2011). Adjusting for bias due to variability of estimated recruitments in fishery assessment models. *Canadian Journal of Fisheries and Aquatic Sciences*, *68*, 1744–1760.

Methot, R. D., & Wetzel, C. R. (2013). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, *142*, 86–99.

Minto, C., Mills Flemming, J., Britten, G. L., & Worm, B. (2014). Productivity dynamics of Atlantic cod. *Canadian Journal of Fisheries and Aquatic Sciences*, *71*, 203–216.

Moustahfid, H., Hendrickson, L. C., Arkhipkin, A., Pierce, G. J., Gangopadhyay, A., Kidokoro, H., Markaida, U., Nigmatullin, C., Sauer, W. H., Jereb, P., Pecl, G., de la Chesnais, T., Ceriola, L., Lazar, N., Firmin, C. J., & Laptikhovsky, V. (2021). Ecological-fishery forecasting of squid Stock dynamics under climate variability and change: Review, challenges, and recommendations. *Reviews in Fisheries Science & Aquaculture*, *29*, 682–705.

Munch, S. B., Giron-Nava, A., & Sugihara, G. (2018). Nonlinear dynamics and noise in fisheries recruitment: A global meta-analysis. *Fish and Fisheries*, *19*, 964–973.

Myers, R. A. (1998). When do environment–recruitment correlations work? *Reviews in Fish Biology and Fisheries*, *8*, 285–305.

Neveu, E., Moore, A. M., Edwards, C. A., Fiechter, J., Drake, P., Crawford, W. J., Jacox, M. G., & Nuss, E. (2016). An historical analysis of the California current circulation using ROMS 4D-Var: System configuration and diagnostics. *Ocean Modelling*, *99*, 133–151.

Ovando, D., Cunningham, C., Kuriyama, P., Boatright, C., & Hilborn, R. (2022). Improving forecasts of sockeye salmon (*Oncorhynchus nerka*) with parametric and nonparametric models. *Canadian Journal of Fisheries and Aquatic Sciences*, *79*, 1198–1210.

Pedersen, E. J., Miller, D. L., Simpson, G. L., & Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ*, *7*, e6876.

Pennekamp, F., Adamson, M. W., Petchey, O. L., Poggiale, J.-C., Aguiar, M., Kooi, B. W., Botkin, D. B., & DeAngelis, D. L. (2017). The practice of prediction: What can ecologists learn from applied, ecology-related fields? *Ecological Complexity*, *32*, 156–167.

Pepin, P., King, J., Holt, C., Gurney-Smith, H., Shackell, N., Hedges, K., & Bundy, A. (2020). Incorporating climate, oceanographic and ecological change considerations into population assessments: A review of Fisheries and Oceans Canada's science advisory process. *Fish and Fisheries*, *23*, 1332–1346.

R Core Development Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Ralston, S., Field, J. C., & Sakuma, K. M. (2015). Long-term variation in a central California pelagic forage assemblage. *Journal of Marine Systems*, *146*, 26–37.

Santora, J. A., Schroeder, I. D., Bograd, S. J., Chavez, F. P., Cimino, M. A., Fiechter, J., Hazen, E. L., Kavanaugh, M. T., Messié, M., Miller, R. R., & Sakuma, K. M. (2021). Pelagic biodiversity, ecosystem function, and services. *Oceanography*, *34*(2), 16–37.

Schaub, M., & Abadi, F. (2011). Integrated population models: A novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology*, *152*, 227–237.

Schroeder, I. D., Santora, J. A., Bograd, S. J., Hazen, E. L., Sakuma, K. M., Moore, A. M., Edwards, C. A., Wells, B. K., & 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. *Canadian Journal of Fisheries and Aquatic Sciences*, *76*, 950–960.

Simon, N., Friedman, J., Tibshirani, R., & Hastie, T. (2011). Regularization paths for cox's proportional hazards model via coordinate descent. *Journal of Statistical Software*, *39*(5), 1–13. https://doi.org/10.18637/jss.v039.i05

Southward, A. J., Langmead, O., Hardman-Mountford, N. J., Aiken, J., Boalch, G. T., Dando, P. R., Genner, M. J., Joint, I., Kendall, M. A., Halliday, N. C., Harris, R. P., Leaper, R., Mieszkowska, N., Pingree, R. D., Richardson, A. J., Sims, D. W., Smith, T., Walne, A. W., & Hawkins, S. J. (2004). Long-term oceanographic and ecological research in the Western English Channel. In A. J. Southward, P. A. Tyler, C. M. Young, & L. A. Fuiman (Eds.), *Advances in marine biology* (pp. 1–105). Academic Press.

Stachura, M. M., Essington, T. E., Mantua, N. J., Hollowed, A. B., Haltuch, M. A., Spencer, P. D., Branch, T. A., & Doyle, M. J. (2014). Linking Northeast Pacific recruitment synchrony to environmental variability. *Fisheries Oceanography*, *23*, 389–408.

Stephens, A., & Taylor, I. G. (2013). *Stock assessment and status of Longspine thornyhead (Sebastolobus altivelis) off California, Oregon, and Washington in 2013*. Pacific Fishery Management Council.

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

Stock, B. C., & Miller, T. J. (2021). The woods hole assessment model (WHAM): A general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates. *Fisheries Research*, *240*, 105967.

Subbey, S., Devine, J. A., Schaarschmidt, U., & Nash, R. D. M. (2014). Modelling and forecasting stock–recruitment: Current and future perspectives. *ICES Journal of Marine Science*, *71*, 2307–2322.

Swalethorp, R., Landry, M., Semmens, B. X., Ohman, M., Aluwihare, L., Chargualaf, D., & Thompson, A. (2022). Anchovy booms and busts linked to trophic shifts in larval diet. *Nature Communications*, *14*, 7412.

Szuwalski, C. S., Vert-Pre, K. A., Punt, A. E., Branch, T. A., & Hilborn, R. (2015). Examining common assumptions about recruitment: A meta-analysis of recruitment dynamics for worldwide marine fisheries. *Fish and Fisheries*, *16*, 633–648.

Thorson, J. T., Stewart, I. J., Taylor, I. G., & Punt, A. E. (2013). Using a recruitment-linked multispecies stock assessment model to estimate common trends in recruitment for US West Coast groundfishes. *Marine Ecology Progress Series*, *483*, 245–256.

Tolimieri, N., & Haltuch, M. A. (2023). Sea-level index of recruitment variability improves assessment model performance for sablefish *Anoplopoma fimbria*. *Canadian Journal of Fisheries and Aquatic Sciences*, *80*, 1006–1016.

Walsh, H. J., Richardson, D. E., Marancik, K. E., & Hare, J. A. (2015). Long-term changes in the distributions of larval and adult fish in the northeast U.S. shelf ecosystem. *PLoS One*, *10*, e0137382.

Walters, C. J., & Collie, J. S. (1988). Is research on environmental factors useful to fisheries management? *Canadian Journal of Fisheries and Aquatic Sciences*, *45*, 1848–1854.

Ward, E. J., Holmes, E. E., Thorson, J. T., & Collen, B. (2014). Complexity is costly: A meta-analysis of parametric and non-parametric methods for short-term population forecasting. *Oikos*, *123*, 652–661.

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

Woolhouse, M. (2011). How to make predictions about future infectious disease risks. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *366*, 2045–2054.

Zimmermann, F., Claireaux, M., & Enberg, K. (2019). Common trends in recruitment dynamics of north-east Atlantic fish stocks and their links to environment, ecology and management. *Fish and Fisheries*, *20*, 518–536.

Zuur, A. F., Fryer, R. J., Jolliffe, I. T., Dekker, R., & Beukema, J. J. (2003). Estimating common trends in multivariate time series using dynamic factor analysis. *Environmetrics*, *14*, 665–685.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ward, E. J., Hunsicker, M. E., Marshall, K. N., Oken, K. L., Semmens, B. X., Field, J. C., Haltuch, M. A., Johnson, K. F., Taylor, I. G., Thompson, A. R., & Tolimieri, N. (2024). Leveraging ecological indicators to improve short term forecasts of fish recruitment. *Fish and Fisheries*, *25*, 895–909. https://doi.org/10.1111/faf.12850