



OPEN

# Semiparametric model selection for identification of environmental covariates related to adult groundfish catches and weights

Hannah E. Correia

Ecologists and fisheries managers are interested in monitoring economically important marine fish species and using this data to inform management strategies. Determining environmental factors that best predict changes in these populations, particularly under rapid climate change, are a priority. I illustrate the application of the least squares-based spline estimation and group LASSO (LSSGLASSO) procedure for selection of coefficient functions in single index varying coefficient models (SIVCMs) on an ecological data set that includes spatiotemporal environmental covariates suspected to play a role in the catches and weights of six groundfish species. Temporal trends in variable selection were apparent, though the selection of variables was largely unrelated to common North Pacific climate indices. These results indicate that the strength of an environmental variable's effect on a groundfish population may change over time, and not necessarily in-step with known low-frequency patterns of ocean-climate variability commonly attributable to large-scale regime shifts in the North Pacific. My application of the LSSGLASSO procedure for SIVCMs to deep water species using environmental data from various sources illustrates how variable selection with a flexible model structure can produce informative inference for remote and hard-to-reach animal populations.

The northern Pacific system is controlled by multiple interdecadal patterns of climate variability that stem from different physical sources<sup>1</sup>. Groundfish populations in the northeastern Pacific Ocean followed the six- to 12-year warming and cooling periods of the El Niño-Southern Oscillation (ENSO)<sup>2</sup>. Sea surface temperature and pressure changes in the North Pacific are captured by the Pacific Decadal Oscillation (PDO), which is separate from ENSO behavior in the region<sup>3</sup>. A third climate cycle described more recently and termed the Northeast Pacific Gyre Oscillation (NPGO) follows variations in ocean nutrient cycling and phytoplankton abundance and plays a role in the larger system of climate variability with ENSO and PDO<sup>4,5</sup>.

Dramatic, permanent changes in marine species compositions in response to shifts in climate modes, commonly referred to as regime shifts, such as the strong one observed in 1976-1977 in the northern Pacific Ocean may be the convergence of several climate patterns switching phases within the same time period<sup>6,7</sup>. This switching of regimes makes it difficult to identify which specific patterns are culprits in affecting distinct marine populations<sup>8,9</sup>. Many studies on fisheries systems continue to focus on these interdecadal climate modes as primary sources of population variability of marine fishes. However, climate modes alone are insufficient to accurately describe variability found in many commercially valuable marine populations on multi-year scales<sup>10</sup>. Other sources of oceanic variability not captured by climate modes exist and should be considered when attempting to create accurate models to describe and predict changes in marine populations, even in areas that appear to be dominated by shifts in climate regimes. The complicated interplay of ocean-climate systems in the North Pacific region makes it difficult to identify which and how specific indices are culprits in affecting distinct marine populations.

Many of the marine fishes in the northern Pacific Ocean are commercially important species that contribute significantly to the economy of the United States and are important sources of food both domestically and internationally<sup>11-13</sup>. Several of these populations are managed by international or regional fishing commissions to control commercial harvests and monitor population health<sup>11,14-16</sup>. These organizations are becoming more concerned about the role climate plays in maintaining healthy fish populations, especially as marine fishes do not recover from population collapses as quickly as previously believed<sup>17</sup>. Fishing activities are increasingly

<sup>1</sup>Department of Biostatistics, Harvard University, Boston, MA 02115, USA. <sup>2</sup>Harvard Data Science Initiative, Harvard University, Cambridge, MA 02138, USA. email: hcorreia@hsph.harvard.edu

concentrated on deeper-dwelling species<sup>18,19</sup>. While focus on the effect of various climate modes has dominated ecological research on fishes in the North Pacific region, relationships between the marine environment and atmospheric trends are nuanced and may involve complex lagged effects, particularly for deepwater populations<sup>20</sup>. Organisms that inhabit the deep ocean are also problematic to study, as they are not adapted to surface-level conditions and prove difficult to sample and keep alive, making experiments in laboratory conditions impossible or prohibitively expensive. Determining which specific environmental variables contribute to fluctuations in the populations of these species from observational data would represent major progress in discerning the impact of climate variability on marine ecosystem health and how those changes affect the economy and food security. A model structure able to accommodate a suite of environmental variables that vary spatiotemporally would be necessary to examine effects of many environmental covariates on deepwater marine populations simultaneously.

Consider the single index varying coefficient model (SIVCM) of the form

$$y_i = \{G(\theta_0^T Z_i)\}^T X_i + \varepsilon_i, \quad i = 1, \dots, n, \quad (1)$$

where  $y_i$  is the response,  $Z = (z_{1i}, z_{2i}, \dots, z_{ti})^T$  and  $X = (x_{0i}, x_{1i}, \dots, x_{pi})^T$  with  $x_{0i} = 1$  are predictor variables,  $\theta_0$  is a vector of unknown coefficients representing the single-index direction,  $G(\cdot) = (g(\cdot)_0, \dots, g(\cdot)_p)^T$  are nonparametric coefficient functions, and  $\varepsilon$  are the random errors<sup>21</sup>. By setting  $Z$  to longitude, latitude, and time triplets, the SIVCM is a convenient structure for incorporating spatiotemporal effects for multiple environmental predictors in  $X$ . Spatiotemporal variation among values is incorporated through three-dimensional functions based on spline smoothing<sup>22–27</sup>, accounting for spatial and temporal autocorrelation and improving prediction and inference<sup>28,29</sup>. Similar structures have been widely used in the related spatially-varying coefficient models<sup>30,31</sup> and temporally-varying coefficient models<sup>32</sup>, including applications in forestry<sup>33</sup>, ecology<sup>34</sup>, and economics<sup>35,36</sup>. Spline-based methods, which are often used to estimate  $G(\cdot)$  in SIVCMs<sup>37,38</sup>, are more robust for spatially correlated data and do not require the spatial variation to be specified by a functional form<sup>39,40</sup>.

For such models, selection of important predictor variables in  $X$  is typically of interest. Forward selection, backwards elimination, and stepwise selection methods are unstable for models with many predictors and even with advancements to the algorithms, these methods are considered sub-optimal for variable selection, particularly for high-dimensional models<sup>41</sup>. Penalty-based regression procedures, such as ridge regression and least absolute shrinkage and selection operator (LASSO) estimation, penalize large regression coefficients to reduce overfitting. LASSO additionally performs variable selection by penalizing small regression coefficients to zero, effectively removing these coefficients from the model<sup>42</sup>. LASSO works particularly well for models with many predictors because it shrinks large coefficients to zero rather than minimizing them, and it is computationally efficient<sup>43</sup>. Group LASSO incorporates information about groupings of variables into the penalty function, which is particularly important for categorical predictor variables<sup>44</sup>. While selection for varying coefficient models (VCMs), a lower-order relative of the SIVCM, have built on both the smoothly clipped absolute deviation (SCAD) and LASSO approaches<sup>45–48</sup>, selection procedures of the single-index direction coefficients or the functions in SIVCMs have primarily used SCAD penalties<sup>49–51</sup>. SCAD procedures are unbiased, but they are sensitive to initial estimation and parameter tuning<sup>48</sup>. LASSO procedures are typically simpler to implement than SCAD, and group LASSO correctly selects important variables for VCMs where the number of dimensions far exceeds the number of observations<sup>52</sup>. In this analysis, I used a combination of least squared-based spline estimation and group LASSO (LSSGLASSO) proposed by Sun et al.<sup>53</sup> to select coefficient functions and estimate the index parameters in a SIVCM of spatiotemporally-varying environmental covariates potentially contributing to changes in groundfish populations in the North Pacific Ocean. With this application, I aimed to establish relevant environmental factors that influence populations of focal groundfish species in this region.

## Methods

Annual surveys of several groundfish species are taken at established locations in the waters along the coast of Alaska by the Alaska Fisheries Science Center (AFSC), a division of the National Oceanic and Atmospheric Administration (NOAA). Catch per unit effort (CPUE), also referred to as catch rate, and mean weight in kilograms of six groundfish species determined at each location for each survey year were obtained for years 1979–2013<sup>54,55</sup>. The six groundfish species of focus in this analysis were Pacific cod (*Gadus macrocephalus*), Pacific halibut (*Hippoglossus stenolepis*), sablefish (*Anoplopoma fimbria*), rougheye rockfish (*Sebastes aleutianus*), shortraker rockfish (*Sebastes borealis*), and shortspine thornyhead (*Sebastolobus alascanus*). Air temperature in degrees Celsius (ATMP), sea level pressure in hPa (PRES), wind speed in meters per second averaged over eight-minute periods (WSPD), sea surface temperature in degrees Celsius (WTMP), and the average height in meters of the highest one-third of all waves in 20-minute sampling periods (WVHT) measured daily from buoys in the Gulf of Alaska were obtained from the National Data Buoy Center and summarized by monthly means<sup>56</sup>. Temperature in degrees Celsius measured at the sea floor (hereafter bottom temperature) was obtained from the AFSC Resource Assessment and Conservation Engineering (RACE) Division's bottom trawl surveys<sup>57</sup>. Zooplankton biomass volume given in number per cubic meter (hereafter plankton) were obtained from the NOAA's Coastal and Oceanic Plankton Ecology, Production, and Observation Database<sup>58</sup>. Alkalinity (Alk), chlorophyll (Chl), nitrate (NO<sub>3</sub>), dissolved oxygen (Oxy), phosphate (Phos), and silicate (Sil) concentrations at depths of 75, 400, and 900 meters were obtained from the NOAA's World Ocean Database<sup>59</sup>.

Since environmental data can include hundreds of predictor variables with potential multi-collinearity, a practical method for reducing variables before performing group LASSO is necessary for the LSSGLASSO procedure to be computationally feasible. Therefore, backward variable elimination of explanatory environmental variables using computed variance inflation factors (VIF) was performed. The VIF for explanatory variable  $i$  was calculated as  $1/(1 - R_i^2)$  where  $R_i^2$  was the correlation coefficient of the linear model with variable  $i$  regressed against all other explanatory variables. VIFs for all explanatory variables were computed, and the variable with

largest VIF that exceeded a value of 5 was removed<sup>60,61</sup>. The process was repeated after each removed variable, until no variables had a  $VIF > 5$ . None of the environmental variables considered for inclusion in the SIVCM had a  $VIF > 5$ . The max VIF for any explanatory variables included in the final model was 2.64.

Measurements of environmental variables were not recorded at the exact same locations across all years for which groundfish surveys were performed, thus spatiotemporal interpolation via inverse distance weighting was used to obtain monthly environmental measures for 1979–2013 at the precise locations where MESA surveys were conducted<sup>62,63</sup>. For all environmental variables, a seasonal amplitude was then calculated for each survey year for all locations. For physical variables ATMP, PRES, WSPD, WTMP, WVHT and bottom temperature, seasonal amplitude was defined as the mean of June, July, and August averages minus the mean of December, January, and February averages. This was owing to the fact that temperatures maximize in the summer and minimize in winter and winter storms enable strong mixing of ocean nutrients<sup>64,65</sup>. Seasonal amplitudes for chemical and biological variables including plankton, Chl, Alk, NO<sub>3</sub>, Oxy, Phos, Sal, and Sil were calculated as the mean of August, September, and October averages minus the mean of March, April, and May averages. Zooplankton biomass increases in May after winter vertical mixing of deepwater nutrients and benefits from spring phytoplankton blooms but has not yet been depleted by grazing from summer-migrating pelagic fishes and cephalopods<sup>66,67</sup>. Therefore nutrients are generally maximal in the spring and minimal in fall after depletion by phytoplankton, whereas zooplankton and chlorophyll would be maximal in fall after nutrient consumption and growth over the summer in the North Pacific region<sup>68–70</sup>.

The AFSC groundfish surveys sample only adults<sup>55</sup>, and groundfish recruitment is related to the effects of physical and biological variables on early life stages<sup>71</sup>. Therefore, all environmental predictors were lagged based on the sexual maturity of the focal groundfish. Sablefish and Pacific cod female maturity is reached around 5 years old<sup>72,73</sup>. Average age of Pacific halibut females that have reached sexual maturity is 12 years, while for males reaching maturity the average age is 8 years<sup>74</sup>. Rockfish sexual maturity is typically attained from 3 to 7 years of age<sup>75</sup>. Therefore, lags of 5 years were used for environmental variables in models with sablefish, rockfish, or Pacific cod CPUEs and weights as responses. Lags of 10 years were used for environmental variables in models with responses of Pacific halibut CPUE and weight. These lags were also supported by the findings of Sun et al.<sup>53</sup>.

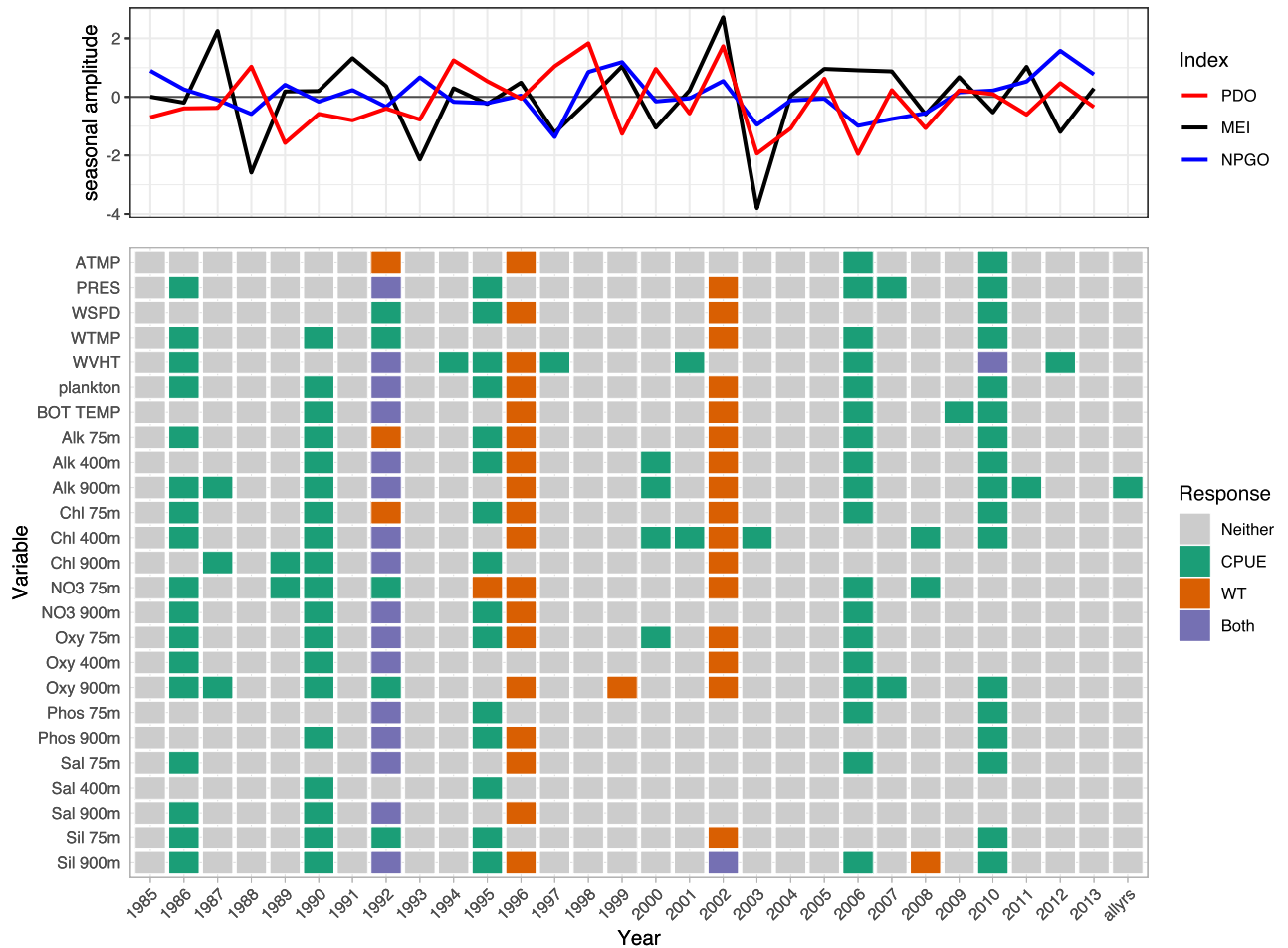
A SIVCM of the form given in (1) was fit, where  $y_i$  was the CPUE or mean weight of a groundfish at each location for each year,  $X = (x_{0i}, \dots, x_{25i})^T$  with  $x_{0i} = 1$  and  $x_{1i}, \dots, x_{25i}$  were the lagged seasonal amplitudes of environmental variables described previously, and  $Z = (z_{1i}, z_{2i}, z_{3i})^T$  were the longitude, latitude, and year for each observation. Selection of important variables was performed per year and over all years using the LSSGLASSO procedure<sup>53</sup>. The choice of knots for the B-spline approximation of the coefficient functions and the optimization of the tuning parameter in the LSSGLASSO procedure are covered in detail in Sun et al.<sup>53</sup> The optimal tuning parameter was chosen using a BIC-type selection criterion, while the number of knots was fixed at eight. The number of knots can be chosen using a selection criterion such as GCV or BIC, but since basis expansion was only used for selection of the coefficient functions and every function was re-estimated using local linear regression, the choice of the number of knots does not affect the results<sup>53</sup>. The performance of the group LASSO procedure was tested against procedures where the group LASSO penalty was replaced by either a group LASSO with ridge penalty or a group SCAD with ridge penalty, and there were no appreciable differences between the three selection procedures. Heat maps were used to visualize selection of variables for each groundfish species. After selection, the SIVCM for all years of sablefish CPUE was refit with scaled  $y$ ,  $X$ , and  $Z$  using only variables selected by LSSGLASSO. The selected coefficient functions for important predictors of sablefish CPUE were plotted to provide a detailed example of interpretation for the SIVCM functions.

To further explore if regional climate conditions influenced the selection of variables each year, I considered potential relationships between the selection of a variable over time and three climate indices that influence environmental systems in the North Pacific region. The Pacific Decadal Oscillation (PDO) monthly index, multivariate El Niño/Southern Oscillation bi-monthly index (MEI), and North Pacific Gyre Oscillation (NPGO) monthly index were obtained for 1979–2013<sup>76–78</sup>. Seasonal amplitude for climate indices was defined for each year as the mean of June, July, and August minus the mean of December, January, and February, since climate indices typically describe physical environmental conditions that contribute to temperature and pressure changes affecting ocean nutrient mixing<sup>79,80</sup>. Logistic regression was then used to fit the model  $g(\mathbb{E}(y_i)) = \beta_0 + \beta_k x_{ki}$  for years  $i = 1985, \dots, 2013$  for each variable within a groundfish response, where  $g(\cdot)$  was a log-link function;  $\beta_0$  was the intercept term;  $\beta_k$  were linear coefficients,  $k \in \{\text{PDO}, \text{MEI}, \text{NPGO}\}$ ;  $x_{ki}$  were seasonal amplitudes of index  $k$  lagged by 5 years when considering the responses of all groundfish except those of Pacific halibut, which were lagged 10 years; and  $y_i$  were binary indicators of whether a variable was or was not selected for each year. A Chi-square test was used to compare the models to an intercept-only model.  $P$ -values from the Chi-square test were adjusted within groundfish response to control for the false discovery rate (FDR) of multiple testing using the Benjamini-Hochberg procedure<sup>81</sup>.

Finally, it was necessary to determine whether PDO, MEI, or NPGO were good predictors of groundfish catches and weights and if the effect of a climate index on the selection of variables was related to that index's direct relationship to the groundfish populations. I fit generalized additive models (GAMs) of the form

$$\mathbb{E}(y_i) = f_0 + f_1(\text{PDO}_i - \text{lag}) + f_2(\text{MEI}_i) + f_3(\text{NPGO}_i), \quad i = 1985, \dots, 2013,$$

with lagged seasonal amplitudes of PDO, MEI, and NPGO as three additive smooth predictors and groundfish CPUEs or weights as responses  $y_i$ <sup>26</sup>. Relationships of climate to groundfish responses cannot be assumed to be linear<sup>82–84</sup>. Therefore, the GAM structure was preferred over a generalized linear model, because it permits nonlinear relationships between predictors and the expected response.  $P$ -values were estimated for each smooth predictor in each model and plotted in a heatmap for efficient visualization.

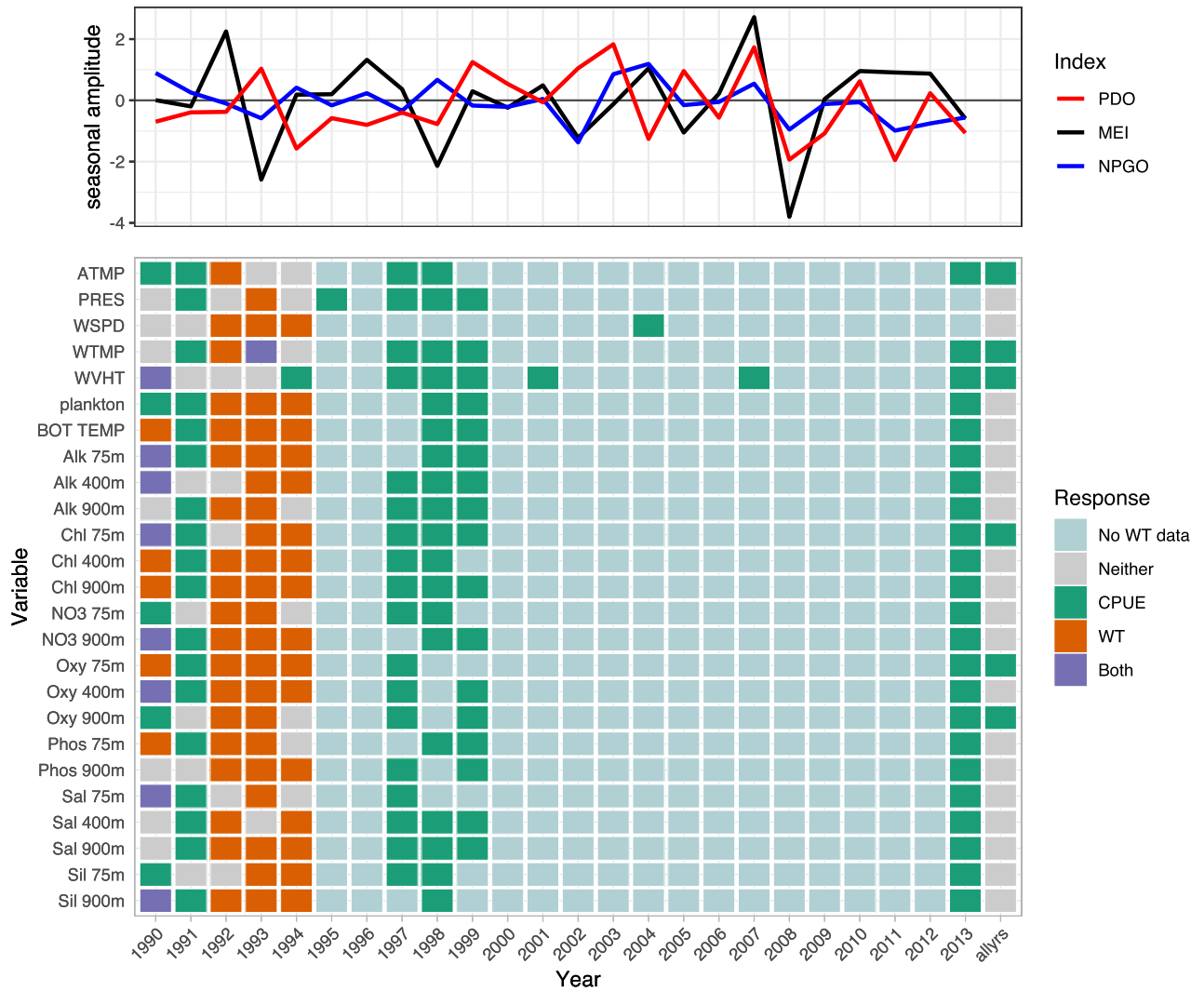


**Figure 1.** Variable selection analyzed for each available year and all years together for the SIVCMs with Pacific cod CPUE or WT as the response. X-axis represents years for which selection was performed, and all available years of data used in selection labeled as “all yrs”. Y-axis contains all variables in SIVCM from which selection was performed. Colors indicate selection of variable(s) important to CPUE, WT, neither response, or both responses, along with years missing either CPUE or WT response values. Line plot at top shows lagged seasonal amplitude values for the Pacific Decadal Oscillation (PDO), multivariate El Niño/Southern Oscillation (MEI), and North Pacific Gyre Oscillation (NPGO) for corresponding years.

### Results

Variables consistently selected by LSSGLASSO as important to Pacific cod CPUE were WVHT, Alk 900m, Chl 400m, Oxy 900m, Sil 900m, PRES, plankton, Alk 400m, NO3 75m, and Oxy 75m, which usually were in years 1986, 1990, 1992, 1995, 2006, and 2010 (Fig. 1, green). Variables consistently selected in 1992, 1996, and 2002 for Pacific cod weight were Sil 900m, WVHT, plankton, bottom temperature, Alk 75m, Alk 400m, Alk 900m, Chl 75m, Chl 400m, NO3 75m, Oxy 75m, and Oxy 900m (Fig. 1, orange). Shared variables selected for both CPUE and weight of Pacific cod were numerous in 1982 (Fig. 1, purple). WVHT, WTMP, ATMP, plankton, Alk 75m, and NO3 900m in years 1990–1991, 1997–1999, and 2013 were selected variables for Pacific halibut CPUE, whereas variables important for Pacific halibut weight included bottom temperature, Alk 75m, Chl 400m, Chl 900m, NO3 900m, Oxy 75m, Oxy 400m, and Sil 900m for years 1990 and 1992–1994 (Fig. 2). ATMP, WTMP, WVHT, Chl 75m, Oxy 75m, and Oxy 900m were important variables for Pacific halibut CPUE when all years of data were considered together. WTMP, WVHT, Alk 900m, and Sil 900m were most often selected as important for sablefish CPUE, while WSPD and ATMP were the only variables selected as important to sablefish weight more than once over the 29-year period of record and were the only variables important to sablefish weight when all years of data were considered together (Fig. 3).

A wide array of environmental variables were selected for rougheye rockfish CPUE consistently and almost exclusively in years 1991, 2001, 2011, and 2013, whereas 1990 was the only year when more than one variable was selected for rougheye rockfish weight (Fig. 4). The most often selected variables related to rougheye rockfish CPUE were WSPD, Oxy 900m, Sil 900m, plankton, Chl 400m, and Oxy 400m, while Chl 400m, Sil 900m, and Oxy 900m were important when all years were considered together. In the case of shorttraker rockfish, most environmental variables were selected as important for both CPUE and weight in 2000 and 2007 (Fig. 5, purple). WSPD, Chl 400m, Oxy 400m, and Oxy 900m were most often selected over time for shorttraker rockfish CPUE, while 900m depths of NO3, Sil, Phos, and Oxy, 75m depths of Alk, Chl, NO3, Oxy, and Sal, and bottom temperature

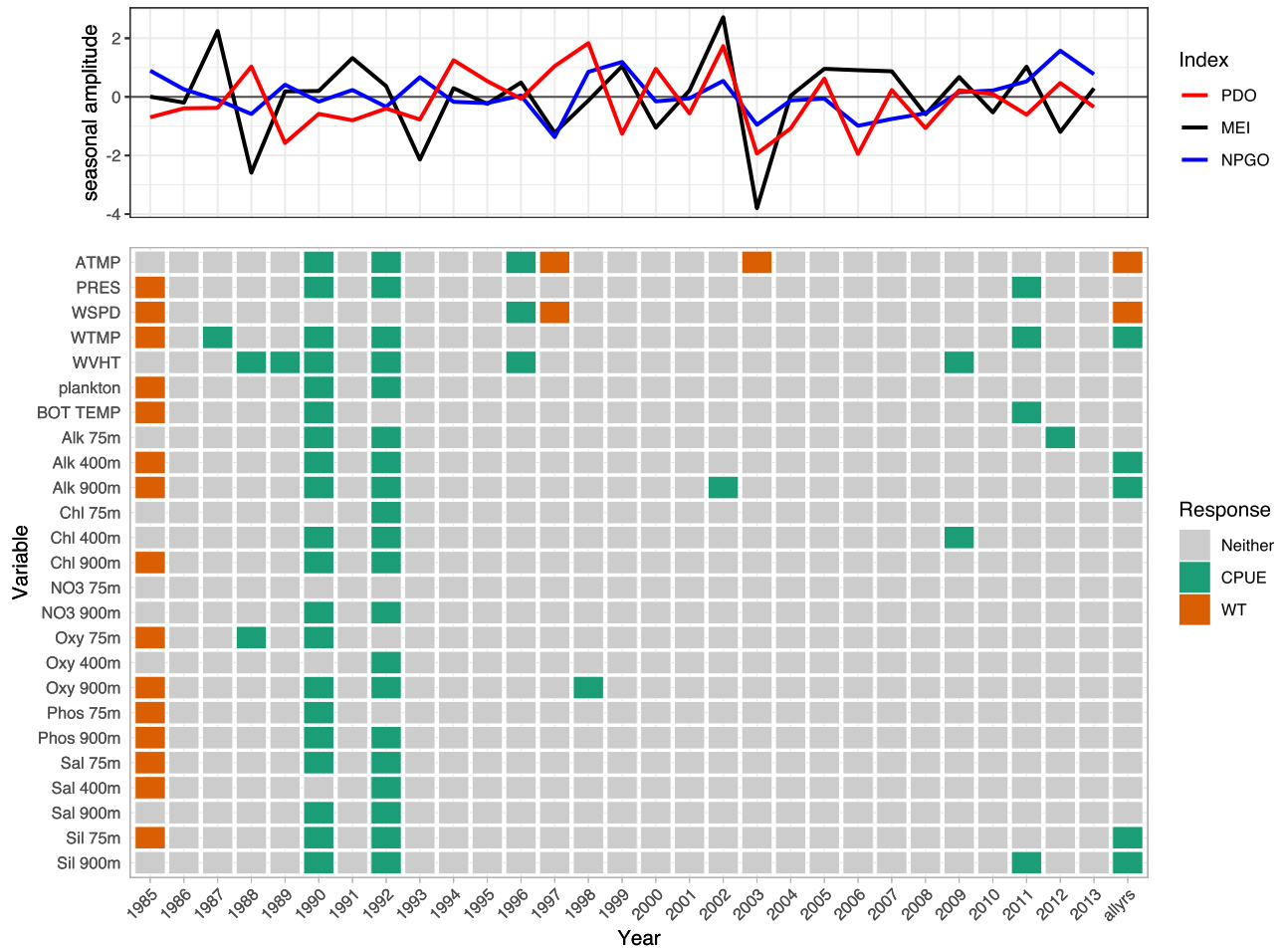


**Figure 2.** Variable selection analyzed for each available year and all years together for the SIVCMs with Pacific halibut CPUE or WT as the response. Axes, colors, and line plot are as described in Fig. 1.

were most often selected as important to shorttraker rockfish weight over all 29 years (Fig. 5). ATMP, WVHT, and Phos 75m were shared as selected variables for shortspine thornyhead CPUE and weight, and many environmental variables contributed to shortspine thornyhead weight in 1986–1987, 1998, 2000, and 2004 (Fig. 6). ATMP was selected most often for shortspine thornyhead CPUE, while WVHT, Chl 400m, and Phos 900m were consistently selected as important to shortspine thornyhead weight over the 29-year period.

Only Pacific cod CPUE, Pacific halibut CPUE, sablefish CPUE and weight, and rougheye rockfish CPUE had variables selected as important predictors when performing selection on data including all available years (Figs. 1, 2, 3, 4). Alk 900m was a common selected variable for Pacific cod and sablefish CPUEs for all years included in the selection procedure, while WTMP was selected for both sablefish and Pacific halibut CPUEs for all years.

I plotted the selected coefficient functions of WTMP, Alk 400m, Alk 900m, Sil 75m, and Sil 900m for sablefish CPUE using all years of data to provide a detailed example of interpretation of the SIVCM coefficient functions (Fig. 7). Parameters representing the single-index direction estimated by the LSSGLASSO procedure were  $\hat{\theta}^T = (0.9346, -0.3380, -0.1109)$ . The relationship between sablefish CPUE and space-time is highly nonlinear, as seen in the intercept function ( $g_0$ ). For a given location and year, we can express the single-index direction as  $\theta^T Z = 0.9346z_1 - 0.3380z_2 - 0.1109z_3$ , where  $z_1$  is longitude,  $z_2$  is latitude, and  $z_3$  is the year. For  $\theta^T Z$  to increase, either longitude increases (moving eastward) or latitude decreases (moving southward). As time moves forward, i.e. year increases,  $\theta^T Z$  decreases. Fixing two variables in  $Z$  while moving the third allows us to consider how each important variable in  $X$  affects the response as  $\theta^T Z$  changes. For example, for a fixed latitude and year, increasing longitude leads to increasing  $\theta^T Z$ , which was associated with increasing, decreasing, and increasing effect of WTMP on sablefish CPUE ( $g_1$  in Fig. 7). So for low and high (east and west) longitude values within the range of the data, increasing WTMP was associated with increasing sablefish CPUE for fixed latitude and year. For medium longitude values, increasing WTMP was accompanied by decreasing sablefish CPUE, ceteris paribus. This same nonlinear relationship was also observed for Sil 900m on sablefish CPUE (Fig. 7,  $g_5$ ). Alternatively for



**Figure 3.** Variable selection analyzed for each available year and all years together for the SIVCMs with sablefish CPUE or WT as the response. Axes, colors, and line plot are as described in Fig. 1.

a fixed longitude and year, decreasing latitude (moving southward) precipitates an increase in  $\theta^T Z$ , which was associated with an increasing and then decreasing effect of Alk 400m on sablefish CPUE ( $g_2$  in Fig. 7). Hence for southern locations with fixed longitude and year, increasing Alk 400m was associated with decreasing sablefish CPUE. As another option for interpretation, increasing year while fixing longitude and latitude leads to a decrease in  $\theta^T Z$ . Therefore for a fixed longitude-latitude pair, increasing years was associated with increasing, decreasing, then increasing effects of Alk 900m on sablefish CPUE (Fig. 7,  $g_3$ ). Thus in early and late years (e.g. 1985–1988 and 2010–2013) for a fixed location, increasing Alk 900m was related to increasing sablefish CPUE.

Lagged seasonal amplitude of NPGO was related to the selection or non-selection of PRES, WSPD, BOT TEMP, Alk 75m, Alk 400m, Chl 75m, Chl 900m, NO3 75m, NO3 900m, Oxy 75m, Oxy 900m, Phos 900m, Sal 75m, Sal 400m, Sil 75m, and Sil 900m for shortraker rockfish weight (Fig. 8). Seasonal amplitudes of PDO, MEI, and NPGO climate indices in the North Pacific were not related to the selection of any environmental variables for any other groundfish responses.

Lagged seasonal amplitudes of PDO, MEI, and NPGO were significant predictors of Pacific cod CPUE, while PDO and MEI were not significant predictors for any of the other groundfish responses (Fig. 9). Lagged seasonal amplitudes of PDO, MEI, and NPGO provided a moderate fit for Pacific cod CPUE (deviance explained= 0.581). Lagged NPGO was also a good predictor of rougheye rockfish CPUE and provided a good fit to the data (deviance explained= 0.641).

### Discussion

Deepwater species such as the groundfish examined in this study are difficult to monitor regularly and study extensively, so determining environmental variables that most accurately predict population health of these species is critical. In this study, I illustrated LASSO-type variable selection on the SIVCM capable of including spatiotemporal covariates for modeling large-scale population data, identifying the most significant factors affecting catch and mean weights of groundfish, and determining the functional relationships of these important predictors to the response for inference. Though the environmental variables considered in this application were collected in situ, the SIVCM with group LASSO-type selection would be ideal for spatially or spatiotemporally measured remotely sensed data, such as environmental satellite data.



**Figure 4.** Variable selection analyzed for each available year and all years together for the SIVCMs with rougeye rockfish CPUE or WT as the response. Axes, colors, and line plot are as described in Fig. 1.

CPUE and mean weights of groundfish contribute to calculations of relative population numbers and weights which are used to inform management and policy decisions<sup>85</sup>. The use of relevant environmental covariates influencing groundfish catch and weight can inform spatiotemporal modeling of groundfish distributions and aid prediction of distribution shifts due to climate change<sup>86</sup>. This knowledge is particularly urgent for predicting the effects of climate change and devising adaptive strategies for vulnerable populations such as the three rockfish species included in this study, stocks of which are highly sensitive to climate change in the eastern Bering Sea<sup>87</sup>.

Increased growth rates in juvenile sablefish with warmer water temperatures have been observed in laboratory settings<sup>88,89</sup> and off the coast of Oregon, where early growth was related to recruitment<sup>90</sup>. Increased sablefish recruitment may be reliant on greater stratification developed by warmer surface water temperatures<sup>91</sup>, while warmer deep waters during the sablefish egg stage combined with colder surface waters during the pelagic larvae stage have been related to higher recruitment<sup>92</sup>. The importance of sea surface temperature to sablefish CPUE across all years established in this analysis is therefore in agreement with previous experimental and observational studies and stresses the importance of surface water temperatures on the success of juvenile sablefish<sup>90</sup>. Sogard<sup>90</sup> presumed faster sablefish growth occurred due to high overall productivity reflected in the PDO index, however in this analysis there was no direct relationship between PDO and sablefish CPUE or mean weight, nor was PDO found to be related to the selection of any environmental variables for sablefish CPUE or weight. However, the selection of silicate as important for sablefish CPUE may point to a relationship of primary production on sablefish, as silicate is an important regulator of phytoplankton growth in Alaskan marine waters<sup>93</sup>. The selection of air temperature and wind speed as important across all years for sablefish weight is consistent with the relationships of transport and of July winds during juvenile years to recruitment<sup>92,94</sup> and suggests the importance of air and surface water movements on early-stage sablefish.

Chlorophyll and sea surface temperature were important across all years for Pacific halibut CPUE, along with air temperature, dissolved oxygen, and wave height. These findings are corroborated by previous observations of Pacific halibut. Temperature influences growth on Pacific halibut juveniles, which exhibit compensatory growth

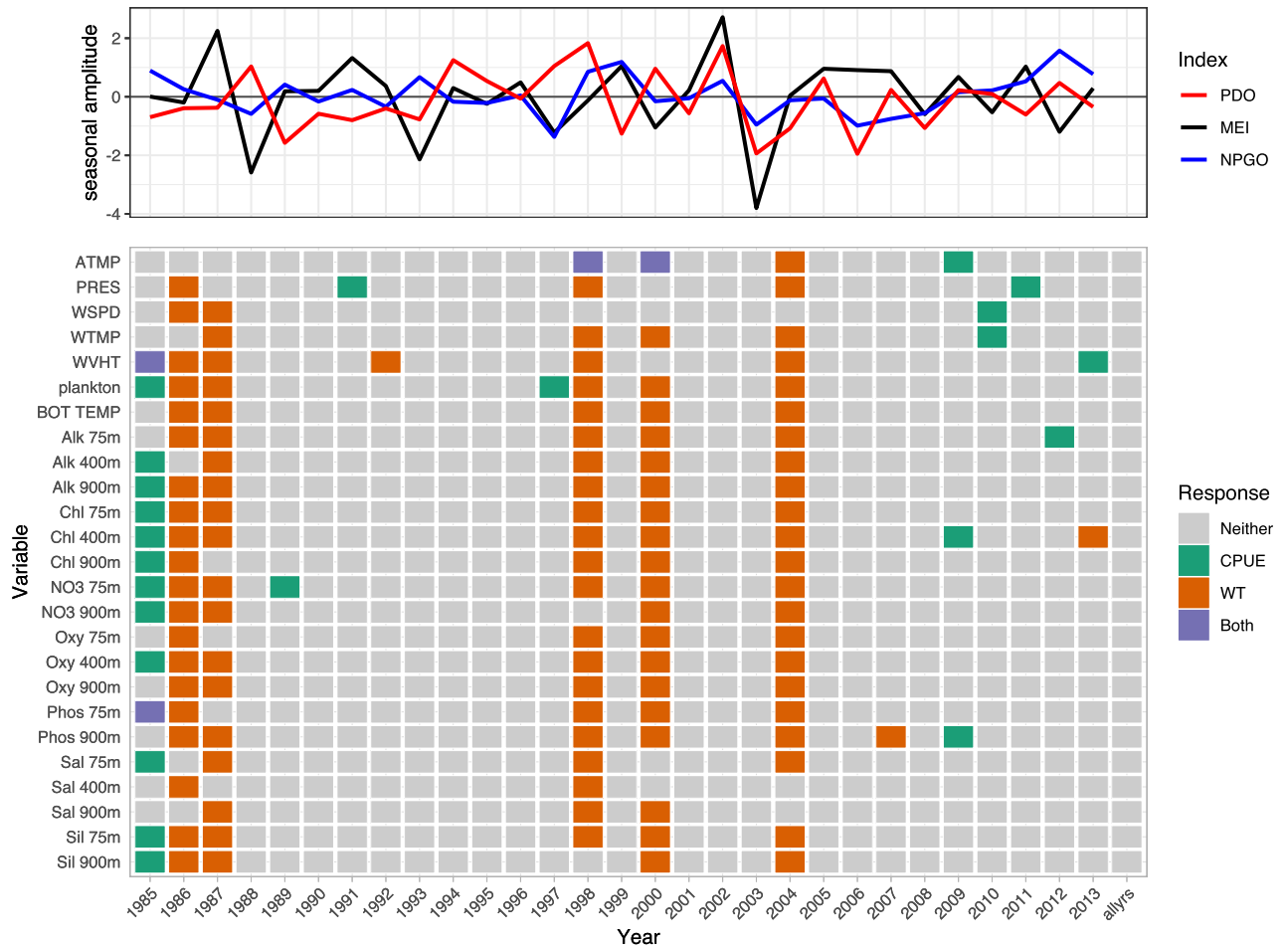


**Figure 5.** Variable selection analyzed for each available year and all years together for the SIVCMs with shortraker rockfish CPUE or WT as the response. Axes, colors, and line plot are as described in Fig. 1.

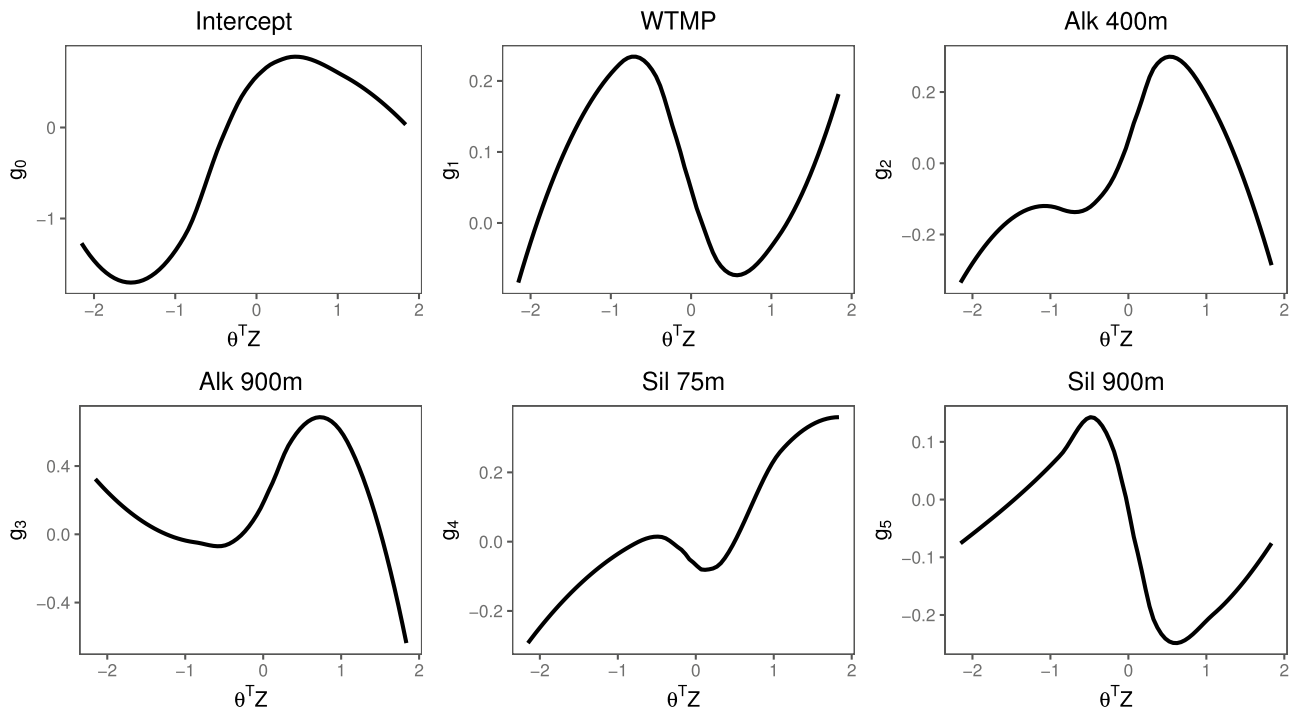
in response to periods of temperature-limited growth<sup>95,96</sup>. Bioenergetics model simulations and diet data suggest that adult growth may also be limited by temperature and prey quality or availability<sup>97</sup>. Pacific halibut weight data were no longer collected after 1994 by the AFSC surveys, so a consistent, long term signal in the data to determine important variables related to weight of Pacific halibut was unlikely. However, growth in early years contributes to year-class abundance for Pacific halibut<sup>95</sup>. The findings of chlorophyll, air temperature, and wave height as important to Pacific halibut CPUE point to sensitivity of Pacific halibut to bottom-up forces in the North Pacific Ocean<sup>98</sup>. Chlorophyll is an important indicator of changes in the carbon transport system and plankton biomass vital for energy transfer to higher trophic levels<sup>69,99</sup>, while air temperature and wave height influence the duration of wind-mixing events and change the depth of the mixed-layer, affecting primary and secondary production<sup>7,100–102</sup>. A relationship between dissolved oxygen and Pacific halibut catch has also been established previously, with Pacific halibut likely adjusting to changes in oxygen environments through lateral rather than vertical movements<sup>103</sup>.

Variables selected as important to specific groundfish in this work that have not previously been explored encourage further investigation. For example, only deep water alkalinity contributed to Pacific cod CPUE across all years of available data, but Pacific cod CPUE was also related to PDO, MEI, and NPGO indices. While water temperatures have been related to interannual depth changes, range shifts, and population declines of adult Pacific cod<sup>104–106</sup>, juvenile Pacific cod may respond to drivers other than temperature<sup>106</sup>. Alkalinity, as part of the ocean's carbonate system, moderates seasonal pH variability<sup>107</sup>. Further, increased variability in the interannual carbonate system in Pacific Ocean is also attributed to variability in the PDO and ENSO<sup>108</sup>. Therefore the selection of alkalinity as important to Pacific cod catch may correspond to several climate modes strongly dictated by the carbonate system and closely correlated to variability in alkalinity. Then again, the importance of alkalinity may instead represent a sensitivity of early stage Pacific cod to ocean acidification, as increased atmospheric carbon dioxide changes alkalinity levels<sup>109,110</sup>.

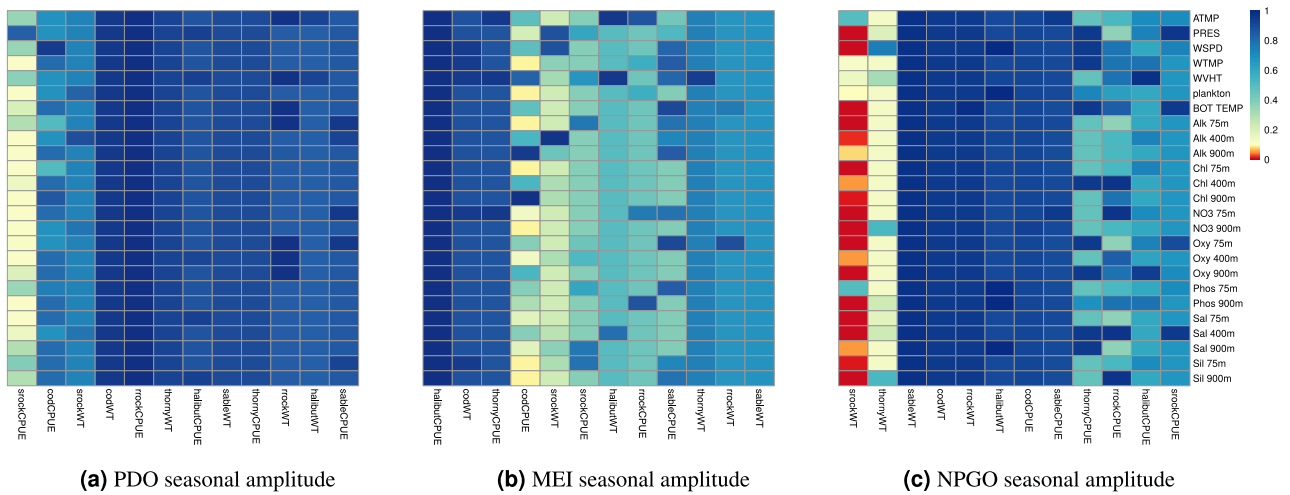




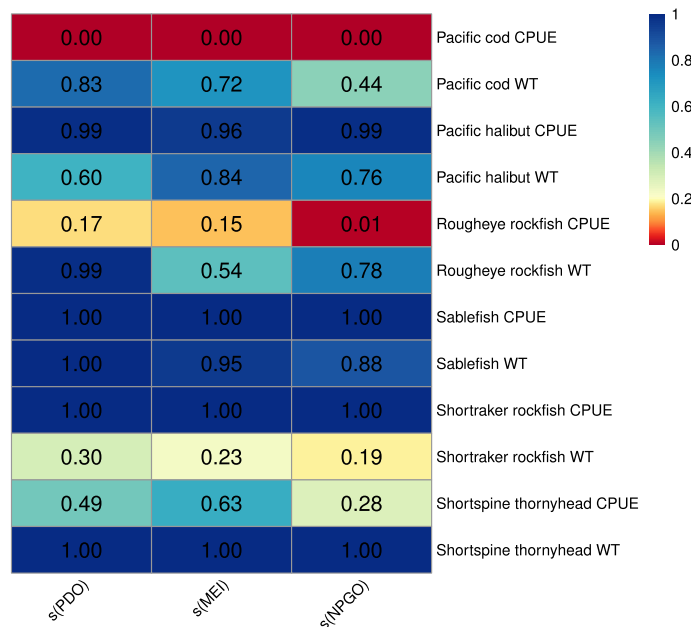
**Figure 6.** Variable selection analyzed for each available year and all years together for the SIVCMs with shortspine thornyhead CPUE or WT as the response. Axes, colors, and line plot are as described in Fig. 1.



**Figure 7.** Coefficient functions selected and estimated by LSSGLASSO as important predictors of sablefish CPUE.



**Figure 8.** FDR-corrected  $P$ -values for relationship between selection of each environmental variable contributing to groundfish CPUEs or WTs each year and lagged seasonal amplitudes of climate indices PDO (a), MEI (b), and NPGO (c) fitted using logistic regression. Abbr: cod = Pacific cod; halibut = Pacific halibut; sable = sablefish; rock = roughey rockfish; srock = shortraker rockfish; thorny = shortspine thornyhead; CPUE = catch per unit effort; WT = mean weight in kg.  $P < 0.05$  are indicated with red.



**Figure 9.** Heatmap of  $P$ -values for nonlinear relationship between groundfish CPUEs or WTs each year and lagged yearly seasonal amplitudes of climate indices PDO, MEI, and NPGO fitted using generalized additive models (GAMs). Colors and abbr. are as described in Fig. 8.

Another example of previously unexplored environmental variables found in this study to be related to groundfish are the important variables for roughey rockfish CPUE across all years, which were exclusively from deeper waters (>300 m). Since rockfish are viviparous, exposure to near-surface environment during development is eliminated. Roughey rockfish CPUE was best predicted by NPGO seasonal amplitudes, while chlorophyll, dissolved oxygen, and silicate were all important for roughey rockfish CPUE over all years. NPGO describes variations in salinity, nitrate, phosphate, silicate, oxygen, and chlorophyll, capturing important interannual and decadal biological patterns<sup>4</sup>. This suggests that the selection of variables affecting roughey rockfish catch was the result of an existing, strong effect of nutrient cycling patterns captured by the NPGO index on roughey rockfish populations.

Selection of important variables related to groundfish catch and weight using LSSGLASSO for SIVCMs provided important confirmation of previous findings and new avenues of research for large populations of groundfish in the North Pacific over an extended time period during which several large-scale regime shifts have been

reported<sup>1,6,111,112</sup>. While large-scale climate indices are often favored when studying populations affected by regime shifts of the northern Pacific marine system, the specific pathways of their effects on a population or community are often difficult to disentangle. Most groundfish responses had a large array of environmental variables selected for specific years, indicating that the strength of environmental effects on groundfish may vary over time. Selection of environmental variables appeared to be unrelated to seasonal amplitudes of the PDO, MEI, and NPGO climate indices for most groundfish. Large-scale climate patterns are associated with groundfish responses<sup>1,2</sup>, however my results highlight the insufficiency of climate modes alone to accurately describe variability found in many commercially valuable marine populations on decadal scales<sup>10</sup>. These results indicate the existence of several pathways of climate and other environmental effects on groundfish responses. Ecological variability in these marine systems are therefore likely driven by gradual spatial and temporal climate variability that reduces resilience, with abrupt changes in climate modes precipitating permanent changes in vulnerable systems<sup>8</sup>.

Received: 11 September 2020; Accepted: 21 April 2021

Published online: 11 May 2021

## References

- Francis, R. C., Hare, S. R., Hollowed, A. B. & Wooster, W. S. Effects of interdecadal climate variability on the oceanic ecosystems of the NE Pacific. *Fish. Oceanogr.* **7**, 1–21. <https://doi.org/10.1046/j.1365-2419.1998.00052.x> (1998).
- Hollowed, A. B. & Wooster, W. S. Variability of winter ocean conditions and strong year classes of Northeast Pacific groundfish. *ICES Mar. Sci. Symp.* **195**, 433–444 (1992).
- Mantua, N. J. & Hare, S. R. The Pacific decadal oscillation. *J. Oceanogr.* **58**, 35–44. <https://doi.org/10.1023/A:1015820616384> (2002).
- Di Lorenzo, E. *et al.* North Pacific gyre oscillation links ocean climate and ecosystem change. *Geophys. Res. Lett.* <https://doi.org/10.1029/2007GL032838> (2008).
- Di Lorenzo, E. *et al.* Synthesis of Pacific ocean climate and ecosystem dynamics. *Oceanography* **26**, 68–81. <https://doi.org/10.5670/oceanog.2013.76> (2013).
- Anderson, P. J. & Piatt, J. F. Community reorganization in the Gulf of Alaska following ocean climate regime shift. *Mar. Ecol. Prog. Ser.* **189**, 117–123 (1999).
- Polovina, J. J., Mitchum, G. T. & Evans, G. T. Decadal and basin-scale variation in mixed layer depth and the impact on biological production in the Central and North Pacific, 1960–88. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **42**, 1701–1716 (1995).
- Litzow, M. A. & Mueter, F. J. Assessing the ecological importance of climate regime shifts: an approach from the North Pacific Ocean. *Prog. Oceanogr.* **120**, 110–119. <https://doi.org/10.1016/j.pocean.2013.08.003> (2014).
- Möllmann, C., Folke, C., Edwards, M. & Conversi, A. Marine regime shifts around the globe: theory, drivers and impacts. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20130260. <https://doi.org/10.1098/rstb.2013.0260> (2015).
- Litzow, M. A., Mueter, F. J. & Hobday, A. J. Reassessing regime shifts in the North Pacific: incremental climate change and commercial fishing are necessary for explaining decadal-scale biological variability. *Glob. Change Biol.* **20**, 38–50. <https://doi.org/10.1111/gcb.12373> (2014).
- Goen, J., & Erikson, L. *Fishery Statistics*. Technical Report. IPHC-2018-AM094-05, International Pacific Halibut Commission (2017).
- Johnson, K. F. *et al.* *Status of the U.S. Sablefish Resource in 2015*. Technical Report. Pacific Fishery Management Council (2016).
- Pacific Fishery Management Council. *Pacific Coast Groundfish Fishery Management Plan*. Technical Report, NOAA (2016).
- NPFMC. *Fishery Management Plan for Groundfish of the Gulf of Alaska*. Technical Report, North Pacific Fishery Management Council (2017).
- Pennoyer, S. & Balsiger, J. *Groundfish Total Allowable Catch Specifications and Prohibited Species Catch Limits Under the Authority of the Fishery Management Plans for the Groundfish Fishery of the Bering Sea and Aleutian Islands Area and Groundfish of the Gulf of Alaska: Final Supplemental Environmental Impact Statement*. Technical Report, United States National Marine Fisheries Service Alaska Regional Office, Juneau, Alaska (1998).
- Rodgveller, C. J., Lunsford, C. R. & Fujioka, J. T. Evidence of hook competition in longline surveys. *Fish. Bull.* **106**, 364–374 (2008).
- Hutchings, J. A. Collapse and recovery of marine fishes. *Nature* **406**, 882–885. <https://doi.org/10.1038/35022565> (2000).
- Moore, J. A. Deep-sea finfish fisheries: lessons from history. *Fisheries* **24**, 16–21 (1999).
- Moore, J. & Mace, P. Challenges and prospects for deep-sea finfish fisheries. *Fisheries* **24**, 22–23 (1999).
- Rijnsdorp, A. D., Peck, M. A., Engelhard, G. H., Möllmann, C. & Pinnegar, J. K. Resolving the effect of climate change on fish populations. *ICES J. Mar. Sci. J. Conseil* **66**, 1570–1583. <https://doi.org/10.1093/icesjms/fsp056> (2009).
- Xia, Y. & Li, W. K. On single-index coefficient regression models. *J. Am. Stat. Assoc.* **94**, 1275–1285. <https://doi.org/10.1080/01621459.1999.10473880> (1999).
- Kammann, E. E. & Wand, M. P. Geoadaptive models. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **52**, 1–18. <https://doi.org/10.1111/1467-9876.00385> (2003).
- Lu, Z., Steinskog, D. J., Tjøstheim, D. & Yao, Q. Adaptively varying-coefficient spatiotemporal models. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **71**, 859–880. <https://doi.org/10.1111/j.1467-9868.2009.00710.x> (2009).
- Ruppert, D., Wand, M. P. & Carroll, R. J. *Semiparametric Regression*. Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, 2003).
- Scheipl, F., Staicu, A.-M. & Greven, S. Functional additive mixed models. *J. Comput. Graph. Stat.* **24**, 477–501. <https://doi.org/10.1080/10618600.2014.901914> (2015).
- Wood, S. N. *Generalized Additive Models*. Texts in Statistical Science Series (Chapman & Hall/CRC, 2006). An Introduction with R.
- Wood, S. N., Scheipl, F. & Faraway, J. J. Straightforward intermediate rank tensor product smoothing in mixed models. *Stat. Comput.* **23**, 341–360. <https://doi.org/10.1007/s11222-012-9314-z> (2013).
- Conn, P. B., Johnson, D. S. & Boveng, P. L. On extrapolating past the range of observed data when making statistical predictions in ecology. *PLoS ONE* **10**, 1–16. <https://doi.org/10.1371/journal.pone.0141416> (2015).
- Hodges, J. S. & Reich, B. J. Adding spatially-correlated errors can mess up the fixed effect you love. *Am. Stat.* **64**, 325–334. <https://doi.org/10.1198/tast.2010.10052> (2010).
- Kim, M. & Wang, L. Generalized spatially varying coefficient models. *J. Comput. Graph. Stat.* <https://doi.org/10.1080/10618600.2020.1754225> (2020).
- Mu, J., Wang, G. & Wang, L. Estimation and inference in spatially varying coefficient models. *Environmetrics* **29**, e2485. <https://doi.org/10.1002/env.2485> (2018).
- Brumback, B. A. & Rice, J. A. Smoothing spline models for the analysis of nested and crossed samples of curves. *J. Am. Stat. Assoc.* **93**, 961–976. <https://doi.org/10.1080/01621459.1998.10473755> (1998).

33. Augustin, N. H., Trenkel, V. M., Wood, S. N. & Lorance, P. Space-time modelling of blue ling for fisheries stock management. *Environmetrics* **24**, 109–119. <https://doi.org/10.1002/env.2196> (2013).
34. Finley, A. O. Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence. *Methods Ecol. Evol.* **2**, 143–154. <https://doi.org/10.1111/j.2041-210X.2010.00060.x> (2011).
35. Al-Sulami, D., Jiang, Z., Lu, Z. & Zhu, J. Estimation for semiparametric nonlinear regression of irregularly located spatial time-series data. *Econom. Stat.* **2**, 22–35. <https://doi.org/10.1016/j.ecosta.2017.01.002> (2017).
36. Gelfand, A. E., Kim, H.-J., Sirmans, C. F. & Banerjee, S. Spatial modeling with spatially varying coefficient processes. *J. Am. Stat. Assoc.* **98**, 387–396. <https://doi.org/10.1198/016214503000170> (2003).
37. Feng, S. & Xue, L. Model detection and estimation for single-index varying coefficient model. *J. Multivariate Anal.* **139**, 227–244. <https://doi.org/10.1016/j.jmva.2015.03.008> (2015).
38. Zhao, P. & Xue, L. Variable selection for semiparametric varying coefficient partially linear models. *Stat. Probab. Lett.* **79**, 2148–2157 (2009).
39. Guisan, A. Jr. & Hastie, T. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol. Model.* **157**, 89–100. [https://doi.org/10.1016/S0304-3800\(02\)00204-1](https://doi.org/10.1016/S0304-3800(02)00204-1) (2002).
40. Zhang, L. & Gove, J. H. Spatial assessment of model errors from four regression techniques. *For. Sci.* **51**, 334–346. <https://doi.org/10.1093/forestscience/51.4.334> (2005).
41. Cai, A., Tsay, R. S. & Chen, R. Variable selection in linear regression with many predictors. *J. Comput. Graph. Stat.* **18**, 573–591 (2009).
42. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **58**, 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x> (1996).
43. Ledolter, J. *Penalty-Based Variable Selection in Regression Models with Many Parameters (LASSO)*, chap. 6, 71–82 (Wiley, 2013). <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118596289.ch6>.
44. Yuan, M. & Lin, Y. Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **68**, 49–67. <https://doi.org/10.1111/j.1467-9868.2005.00532.x> (2006).
45. Fan, J., Yao, Q. & Cai, Z. Adaptive varying-coefficient linear models. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **65**, 57–80 (2003).
46. Matsui, H. & Misumi, T. Variable selection for varying-coefficient models with the sparse regularization. *Comput. Stat.* **30**, 43–55 (2015).
47. Wang, H. & Xia, Y. Shrinkage estimation of the varying coefficient model. *J. Am. Stat. Assoc.* **104**, 747–757. <https://doi.org/10.1198/jasa.2009.0138> (2009).
48. Xue, L. & Qu, A. Variable selection in high-dimensional varying-coefficient models with global optimality. *J. Mach. Learn. Res.* **13**, 1973–1998 (2012).
49. Feng, S. & Xue, L. Variable selection for single-index varying-coefficient model. *Front. Math. China* **8**, 541–565. <https://doi.org/10.1007/s11464-013-0284-z> (2013).
50. Song, Y., Jian, L. & Lin, L. Robust exponential squared loss-based variable selection for high-dimensional single-index varying-coefficient model. *J. Comput. Appl. Math.* **308**, 330–345. <https://doi.org/10.1016/j.cam.2016.05.030> (2016).
51. Yang, J. & Yang, H. Robust modal estimation and variable selection for single-index varying-coefficient models. *Commun. Stat. Simul. Comput.* **46**, 2976–2997. <https://doi.org/10.1080/03610918.2015.1069346> (2017).
52. Wei, F., Huang, J. & Li, H. Variable selection and estimation in high-dimensional varying-coefficient models. *Stat. Sin.* **21**, 1515–1540. <https://doi.org/10.5705/ss.2009.316> (2011).
53. Sun, W., Bindele, H. F., Abebe, A. & Correia, H. E. Robust functional coefficient selection for the single-index varying coefficients regression model. *J. Stat. Comput. Simul.* **1**, 17. <https://doi.org/10.1080/00949655.2020.1867548> (2021).
54. Alaska Fisheries Science Center. AFSC/ABL: Longline Sablefish Survey. <https://noaa-fisheries-afsc.data.socrata.com/dataset/AFSC-ABL-Longline-Sablefish-Survey/itxd-qjvg/data> (2019). Accessed 14 Apr 2014.
55. Sigler, M. F. & Lunsford, C. R. *Survey Protocol for the Alaska Sablefish Longline Survey*. Technical Report, Alaska Fisheries Science Center (2009).
56. National Data Buoy Center. Meteorological and oceanographic data collected from the National Data Buoy Center Coastal-Marine Automated Network (C-MAN) and moored (weather) buoys. <https://accession.nodc.noaa.gov/NDBC-CMANWx> (2018).
57. Alaska Fisheries Science Center. AFSC/RACE/GAP: RACEBASE Database. Online: [http://www.afsc.noaa.gov/RACE/groun dfish/survey\\_data/default.htm](http://www.afsc.noaa.gov/RACE/groun dfish/survey_data/default.htm) (2019).
58. O'Brien, T. D. *COPEPOD: The Global Plankton Database. A Review of the 2007 Database Contents and New Quality Control Methodology*. Technical Report. NOAA Tech. Memo. NMFS-F/ST-34, U.S. Dep. Commerce (2007).
59. Boyer, T. P. *et al.* *World Ocean Database 2013*. Technical Report. National Oceanographic Data Center, Ocean Climate Laboratory, NOAA (2013). <https://doi.org/10.7289/V5NZ85MT>.
60. Neter, J., Kutner, M. H., Nachtsheim, C. J. & Wasserman, W. *Applied Linear Statistical Models*, vol. 4 (Irwin Chicago, 1996).
61. O'Brien, R. M. A caution regarding rules of thumb for variance inflation factors. *Qual. Quant.* **41**, 673–690 (2007).
62. Li, L., Losser, T., Yorke, C. & Piltner, R. Fast inverse distance weighting-based spatiotemporal interpolation: a web-based application of interpolating daily fine particulate matter pm2.5 in the contiguous U.S. using parallel programming and k-d tree. *Int. J. Environ. Res. Public Health* **11**, 9101–9141. <https://doi.org/10.3390/ijerph110909101> (2014).
63. Melo, C. & Melo, O. *geosptdb: Spatio-Temporal Inverse Distance Weighting and Radial Basis Functions with Distance-Based Regression* (2015). R package version 0.5-0.
64. Whitney, F. A. Nutrient variability in the mixed layer of the subarctic Pacific Ocean, 1987–2010. *J. Oceanogr.* **67**, 481–492. <https://doi.org/10.1007/s10872-011-0051-2> (2011).
65. Whitney, F. A., Bograd, S. J. & Ono, T. Nutrient enrichment of the subarctic Pacific Ocean pycnocline. *Geophys. Res. Lett.* **40**, 2200–2205. <https://doi.org/10.1002/grl.50439> (2013).
66. Brodeur, R. D. & Ware, D. M. Long-term variability in zooplankton biomass in the subarctic Pacific Ocean. *Fish. Oceanogr.* **1**, 32–38. <https://doi.org/10.1111/j.1365-2419.1992.tb00023.x> (1992).
67. Chiba, S., Tadokoro, K., Sugisaki, H. & Saino, T. Effects of decadal climate change on zooplankton over the last 50 years in the western subarctic North Pacific. *Glob. Change Biol.* **12**, 907–920. <https://doi.org/10.1111/j.1365-2486.2006.01136.x> (2006).
68. Childers, A. R., Whitley, T. E. & Stockwell, D. A. Seasonal and interannual variability in the distribution of nutrients and chlorophyll a across the Gulf of Alaska shelf: 1998–2000. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **52**, 193–216. <https://doi.org/10.1016/j.dsr2.2004.09.018> (2005). U.S. GLOBEC Biological and Physical Studies of Plankton, Fish and Higher Trophic Level Production, Distribution, and Variability in the Northeast Pacific.
69. Sackmann, B., Mack, L., Logsdon, M. & Perry, M. J. Seasonal and inter-annual variability of SeaWiFS-derived chlorophyll a concentrations in waters off the Washington and Vancouver Island coasts, 1998–2002. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **51**, 945–965 (2004).
70. Wong, C. *et al.* Seasonal cycles of nutrients and dissolved inorganic carbon at high and mid latitudes in the North Pacific Ocean during the Skaugran cruises: determination of new production and nutrient uptake ratios. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **49**, 5317–5338 (2002).
71. Houde, E. D. chap. Recruitment variability. In *Fish Reproductive Biology: Implications for Assessment and Management* (Eds. Jakobsen, T., Fogarty, M. J., Megrey, B. A. & Moksness, E.) (Wiley, 2016).

72. Mason, J. C., Beamish, R. J. & McFarlane, G. A. Sexual maturity, fecundity, spawning, and early life history of sablefish (*Anoplopoma fimbria*) off the Pacific Coast of Canada. *Can. J. Fish. Aquat. Sci.* **40**, 2126–2134. <https://doi.org/10.1139/f83-247> (1983).
73. Stark, J. W. Geographic and seasonal variations in maturation and growth of female Pacific cod (*Gadus macrocephalus*) in the Gulf of Alaska and Bering Sea. *Fish. Bull.* **105**, 396–407 (2007).
74. Clark, W. G., Hare, S. R., Parma, A. M., Sullivan, P. J. & Trumble, R. J. Decadal changes in growth and recruitment of Pacific halibut (*Hippoglossus stenolepis*). *Can. J. Fish. Aquat. Sci.* **56**, 242–252. <https://doi.org/10.1139/f98-163> (1999).
75. Echeverria, T. W. Thirty-four species of California rockfishes: maturity and seasonality of reproduction. *Fish. Bull.* **85**, 229–250 (1987).
76. Di Lorenzo, E. North Pacific Gyre Oscillation (2018). NPGO index.
77. NOAA ESRL Physical Sciences Division. Multivariate ENSO Index Version 2 (MEI.v2) (2019). ENSO index.
78. Mantua, N. J. & JISAU, University of Washington. The Pacific Decadal Oscillation. <http://research.jisau.washington.edu/pdo/> (2016).
79. Di Lorenzo, E. *et al.* Central Pacific El Niño and decadal climate change in the North Pacific Ocean. *Nat. Geosc.* **3**, 762–765. <https://doi.org/10.1038/ngeo984> (2010).
80. Ladd, C. & Stabeno, P. J. Stratification on the Eastern Bering Sea shelf revisited. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **65–70**, 72–83. <https://doi.org/10.1016/j.dsr2.2012.02.009> (2012). Understanding Ecosystem Processes in the Eastern Bering Sea.
81. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodol.)* **57**, 289–300 (1995).
82. Fraimer, A. *et al.* Climate-driven changes in functional biogeography of Arctic marine fish communities. *Proc. Natl. Acad. Sci.* **114**, 12202–12207. <https://doi.org/10.1073/pnas.1706080114> (2017).
83. Hewitt, J. E., Ellis, J. I. & Thrush, S. F. Multiple stressors, nonlinear effects and the implications of climate change impacts on marine coastal ecosystems. *Glob. Change Biol.* **22**, 2665–2675. <https://doi.org/10.1111/gcb.13176> (2016).
84. Liu, H. *et al.* Nonlinear dynamic features and co-predictability of the Georges Bank fish community. *Mar. Ecol. Prog. Ser.* **464**, 195–207 (2012).
85. Echave, K., Rodgveller, C. & Shotwell, S. K. *Calculation of the Geographic Area Sizes Used to Create Population Indices for the Alaska Fisheries Science Center Longline Survey*. Technical Report. NOAA Tech. Memo. NMFS-AFSC-253, U.S. Department of Commerce (2013).
86. Webster, R. A., Soderlund, E., Dykstra, C. L. & Stewart, I. J. Monitoring change in a dynamic environment: spatio-temporal modelling of calibrated data from different types of fisheries surveys of Pacific halibut. *Can. J. Fish. Aquat. Sci.* <https://doi.org/10.1139/cjfas-2019-0240> (2020).
87. Spencer, P. D., Hollowed, A. B., Sigler, M. F., Hermann, A. J. & Nelson, M. W. Trait-based climate vulnerability assessments in data-rich systems: an application to eastern Bering Sea fish and invertebrate stocks. *Glob. Change Biol.* **25**, 3954–3971. <https://doi.org/10.1111/gcb.14763> (2019).
88. Sogard, S. M. & Olla, B. L. Growth and behavioral responses to elevated temperatures by juvenile sablefish *Anoplopoma fimbria* and the interactive role of food availability. *Mar. Ecol. Prog. Ser.* **217**, 121–134 (2001).
89. Stoner, A. W. & Sturm, E. A. Temperature and hunger mediate sablefish (*Anoplopoma fimbria*) feeding motivation: implications for stock assessment. *Can. J. Fish. Aquat. Sci.* **61**, 238–246. <https://doi.org/10.1139/f03-170> (2004).
90. Sogard, S. Interannual variability in growth rates of early juvenile sablefish and the role of environmental factors. *Bull. Mar. Sci.* <https://doi.org/10.5343/bms.2010.1045> (2011).
91. Shotwell, S. K., Hanselman, D. H. & Belkin, I. M. Toward biophysical synergy: investigating advection along the Polar Front to identify factors influencing Alaska sablefish recruitment. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **107**, 40–53. <https://doi.org/10.1016/j.dsr2.2012.08.024> (2014). Fronts, Fish and Top Predators.
92. Tolimieri, N., Haltuch, M. A., Lee, Q., Jacox, M. G. & Bograd, S. J. Oceanographic drivers of sablefish recruitment in the California current. *Fish. Oceanogr.* **27**, 458–474. <https://doi.org/10.1111/fog.12266> (2018).
93. Harrison, P. J., Whitney, F. A., Tsuda, A., Saito, H. & Tadokoro, K. Nutrient and plankton dynamics in the NE and NW gyres of the subarctic Pacific Ocean. *J. Oceanogr.* **60**, 93–117 (2004).
94. Coffin, B. & Mueter, F. Environmental covariates of sablefish (*Anoplopoma fimbria*) and Pacific ocean perch (*Sebastes alutus*) recruitment in the Gulf of Alaska. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **132**, 194–209. <https://doi.org/10.1016/j.dsr2.2015.02.016> (2016). Understanding Ecosystem Processes in the Gulf of Alaska: Volume 1.
95. Hagen, P. T. & Quinn, T. J. Long-term growth dynamics of young Pacific halibut: evidence of temperature-induced variation. *Fish. Res.* **11**, 283–306. [https://doi.org/10.1016/0165-7836\(91\)90006-2](https://doi.org/10.1016/0165-7836(91)90006-2) (1991). Fish Population Dynamics: Solving Fishery Management Problems.
96. Hurst, T. P., Spencer, M. L., Sogard, S. M. & Stoner, A. W. Compensatory growth, energy storage and behavior of juvenile Pacific halibut *Hippoglossus stenolepis* following thermally induced growth reduction. *Mar. Ecol. Prog. Ser.* **293**, 233–240 (2005).
97. Holsman, K. K., Aydin, K., Sullivan, J., Hurst, T. & Kruse, G. H. Climate effects and bottom-up controls on growth and size-at-age of Pacific halibut (*Hippoglossus stenolepis*) in Alaska (USA). *Fish. Oceanogr.* **28**, 345–358. <https://doi.org/10.1111/fog.12416> (2019).
98. Lynam, C. P. *et al.* Interaction between top-down and bottom-up control in marine food webs. *Proc. Natl. Acad. Sci.* **114**, 1952–1957. <https://doi.org/10.1073/pnas.1621037114> (2017).
99. Desmit, X., Ruddick, K. & Lacroix, G. Salinity predicts the distribution of chlorophyll a spring peak in the southern North Sea continental waters. *J. Sea Res.* **103**, 59–74. <https://doi.org/10.1016/j.seares.2015.02.007> (2015).
100. Benson, A. J. & Trites, A. W. Ecological effects of regime shifts in the Bering Sea and eastern North Pacific Ocean. *Fish. Fish.* **3**, 95–113. <https://doi.org/10.1046/j.1467-2979.2002.00078.x> (2002).
101. Feng, J. *et al.* Contrasting correlation patterns between environmental factors and chlorophyll levels in the global ocean. *Glob. Biogeochem. Cycles* **29**, 2095–2107. <https://doi.org/10.1002/2015GB005216> (2015).
102. Kahru, M. *et al.* Global correlations between winds and ocean chlorophyll. *J. Geophys. Res. Oceans* <https://doi.org/10.1029/2010JC006500> (2010).
103. Sadorus, L. L., Mantua, N. J., Essington, T., Hickey, B. & Hare, S. Distribution patterns of Pacific halibut (*Hippoglossus stenolepis*) in relation to environmental variables along the continental shelf waters of the US West Coast and southern British Columbia. *Fish. Oceanogr.* **23**, 225–241. <https://doi.org/10.1111/fog.12057> (2014).
104. Barbeaux, S. *et al.* *Gulf of Alaska Stock Assessments*. Technical Report, North Pacific Fishery Management Council, Anchorage, AK (2018).
105. Barbeaux, S. J. & Hollowed, A. B. Ontogeny matters: climate variability and effects on fish distribution in the eastern Bering Sea. *Fish. Oceanogr.* **27**, 1–15. <https://doi.org/10.1111/fog.12229> (2018).
106. Yang, Q. *et al.* How “The Blob” affected groundfish distributions in the Gulf of Alaska. *Fish. Oceanogr.* **28**, 434–453. <https://doi.org/10.1111/fog.12422> (2019).
107. Hagens, M. & Middelburg, J. J. Attributing seasonal pH variability in surface ocean waters to governing factors. *Geophys. Res. Lett.* **43**, 12528–12537. <https://doi.org/10.1002/2016GL071719> (2016).
108. Fry, C. H., Tyrrell, T., Hain, M. P., Bates, N. R. & Achterberg, E. P. Analysis of global surface ocean alkalinity to determine controlling processes. *Mar. Chem.* **174**, 46–57 (2015).

109. Bromhead, D. *et al.* The potential impact of ocean acidification upon eggs and larvae of yellowfin tuna (*Thunnus albacares*). *Deep Sea Res. Part II Top. Stud. Oceanogr.* **113**, 268–279, <https://doi.org/10.1016/j.dsr2.2014.03.019> (2015). Impacts of climate on marine top predators.
110. Doney, S. C. *et al.* Impact of anthropogenic atmospheric nitrogen and sulfur deposition on ocean acidification and the inorganic carbon system. *Proc. Natl. Acad. Sci.* **104**, 14580–14585. <https://doi.org/10.1073/pnas.0702218104> (2007).
111. Napp, J. M. & Hunt, G. L. Anomalous conditions in the south-eastern Bering Sea 1997: linkages among climate, weather, ocean, and biology. *Fish. Oceanogr.* **10**, 61–68. <https://doi.org/10.1046/j.1365-2419.2001.00155.x> (2001).
112. Noakes, D. J. & Beamish, R. J. Synchrony of marine fish catches and climate and ocean regime shifts in the North Pacific Ocean. *Mar. Coast. Fish.* **1**, 155–168. <https://doi.org/10.1577/C08-001.1> (2009).

## Acknowledgements

This material is based upon work supported by the NSF Graduate Research Fellowship under Grant No. DGE-1414475.

## Author contributions

H.E.C. conceived and conducted the analyses, analysed the results, and wrote the manuscript.

## Competing interests

The author declares no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to H.E.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021