LIMNOLOGY
and
OCEANOGRAPHY

ASLO

Limnol. Oceanogr. 64, 2019, 2632–2645

doi: 10.1002/lno.11238

Measuring the impact of oceanographic indices on species distribution shifts: The spatially varying effect of cold-pool extent in the eastern Bering Sea

James T. Thorson *

Habitat and Ecosystem Process Research program, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, Washington

Abstract

Oceanographers have spent decades developing annual indices that summarize physical conditions in marine ecosystems. Examples include the Pacific Decadal Oscillation, summarizing annual variation in the location of warm waters in the North Pacific, and cold-pool extent (CPE), summarizing the area with cold near-bottom waters in the eastern Bering Sea. However, these indices are rarely included in the species distribution models that are used to identify and forecast distribution shifts under future climate scenarios. I therefore review three interpretations of spatially varying coefficient models, explain how they can be used to estimate spatial patterns of population density associated with oceanographic indices, and add this option to the multivariate spatiotemporal model *VAST*. I then use a case study involving bottom trawl data for 17 fish and decapod species in the eastern Bering Sea 1982–2017 to answer: does a spatially varying coefficient model for CPE explain variation in spatial distribution for species in this region? And (2) does a spatially varying effect of CPE remain substantial even when local temperature is also included as a covariate? Results show that CPE and local bottom temperature are both identified as parsimonious by Akaike Information Criterion for 13 of 17 species, jointly explain nearly 9%–14% of spatiotemporal variation on average, and CPE does explain variation in excess of local temperature alone. I therefore conclude that spatially varying coefficient models are a useful way to assimilate oceanographic indices within species distribution models, and hypothesize that these will be useful to account for decadal-scale variability within multidecadal forecasts of distribution shift.

Ecosystem-based management (EBM) involves regulating multiple ocean impacts including harvest, tourism, and energy development while accounting for species interactions, ecosystem drivers, and socioeconomic linkages. EBM has evolved in tandem with new tools for understanding the multiple human impacts and outcomes resulting from ocean management. Examples of new tools include techniques to include annual oceanographic conditions in the stock-assessment models that are used to define annual catch limits (Schirripa et al. 2009), ecosystem models that include mechanistic detail regarding terrestrial and physical drivers (Fulton et al. 2011), and spatial models that are used to estimate and validate maps of fish habitats (Rooper et al. 2016). Continued improvement in these tools is likely to support ongoing developments in ocean management and governance, for example, the Bering Sea Fisheries Ecosystem Plan, which now includes explicit use of climate-linked ecosystem models (North Pacific Fishery Management Council 2019).

One topic of growing importance in ocean governance is climate-driven shifts in species distribution (Pinsky et al. 2018; Karp et al. 2019). Distribution shifts are increasingly identified from multiple data sets using species distribution/density models (SDMs) fitted to occurrence, count, or biomass-sampling data (e.g., Dolder et al. 2018). SDMs have been used extensively in ecology, oceanography, and fisheries science to describe the spatial distribution and ecological niche of marine and terrestrial species worldwide. Interest in SDMs has increased as researchers have sought to train them using historical and contemporary data and then forecast changes in spatial distribution under alternative climate scenarios (Araújo and New 2007). SDMs have been built to predict local density using local environmental conditions and/or via the interaction of annual covariates and spatial coordinates, and size- or age-structured SDMs have also been used to account for size-based processes affecting distribution shifts (Kristensen et al. 2014; Thorson et al. 2015; Kai et al. 2017).

*Correspondence: james.thorson@noaa.gov

Additional Supporting Information may be found in the online version of this article.

However, studies examining forecasted distribution shifts that are compared with subsequent observations have shown that forecast skill is sometimes poor using the current generation of SDMs that are fitted to localized effects of environmental conditions (Thorson 2019*a*). For example, ontogenetic habitat preferences combined with changes in size/age-structure have been hypothesized to drive observed distribution shifts (Barbeaux and Hollowed 2018), but this has not been successful at explaining a large portion of historical distribution shifts for *Gadus chalcogrammus* (Thorson et al. 2017) or *Paralichthys dentatus* (Perretti and Thorson 2019). This poor skill would arise if species distributions are affected by more than local environmental conditions, for example, due to impacts of local predator densities, lagged environmental effects, or geographically distant environmental conditions that affect the preference of mobile species for local habitats. There is therefore an ongoing need to identify improved techniques to forecast distribution shifts.

In particular, multidecadal shifts in distribution are typically forecasted by fitting to local environmental conditions (Pinsky et al. 2018), some assumed relationship with local conditions (Cheung et al. 2008), or theoretical predictions of metabolic constraints (Teal et al. 2018). Although these local environmental conditions may capture long-term changes in the fundamental niche, it is likely that decadal-scale oscillations in oceanographic conditions will also contribute to interannual variation around long-term trends. Physical oceanographers have spent decades developing oceanographic indices that are an integrated and high-level summary of interannual variation in ocean conditions (Grimmer 1963; Kidson 1975), and fisheries scientists have shown that these oceanographic indices can be informative about patterns in fish productivity (O'Leary et al. 2018). For example, the Pacific Decadal Oscillation (PDO) is an annual index of the location of elevated ocean temperatures in the North Pacific relative to their climatological average, and the PDO was originally identified to correlate with oscillating productivity of salmon stocks between the US West Coast and Alaska (Mantua and Hare 2002).

Despite the well-documented role of oceanographic indices in driving changes in fish productivity and distribution, there is surprisingly little research regarding their potential role in SDMs used to forecast marine distribution shifts. As I will show, this may arise due to the way in which oceanographic indices would typically be included in regression models such as SDM, wherein an annual index can only impact spatial distribution through its interaction with a spatially referenced variable. I therefore describe a novel approach to including oceanographic indices via a "spatially varying coefficient" (SVC) model. Instead of estimating a single slope parameter representing the effect of an oceanographic index on density, it estimates a separate slope parameter for every modeled location, and thereby estimates that some locations have increased density given a positive index phase (positive slope) and other locations have decreased density (negative slope). These spatially varying slopes can then be plotted to visualize the expected shifts in distribution associated with positive values of the annual index.

I first discuss several potential interpretations of this SVC model and modify an existing R package *VAST* for multivariate spatiotemporal models (Thorson and Barnett 2017) to include SVC models. I then explore SVC models using a case study involving 17 fish and decapod species in the eastern Bering Sea, where spatial dynamics are widely hypothesized to be linked to the spatial extent of cold waters resulting from ice melt in previous years. I specifically use this case study to answer two questions: (1) does an SVC model for cold-pool extent explain a substantial portion of variation in spatial distribution for species in this region? And (2) does a spatially varying impact of the cold pool remain useful and significant even when also including local temperature as a covariate? I end by discussing the conditions under which an SVC model is expected to explain distribution shifts better than local environmental conditions in isolation.

Methods

I seek to determine whether spatially varying responses to annual oceanographic indices can improve model parsimony, reduce unexplained variation, and decrease forecast errors relative to a model without covariates (whether oceanographic indices or local environmental conditions) and whether SVCs are still estimable and parsimonious even when including local environmental conditions in addition to annual oceanographic indices. To do so, I first introduce SVC models, their interpretation, and their estimation. I then describe a case study involving 17 fish and decapod species in the eastern Bering Sea.

SVC models

One alternative to using local covariate effects is the SVC models. These models have been discussed extensively in statistical applications (Gelfand et al. 2004; Finley 2011) and have seen some limited use in fisheries (Bacheler et al. 2009; Bartolino et al. 2011), for example, to identify seasonal shifts in pollock distribution in the eastern Bering Sea (Bacheler et al. 2012). As a simple example, SVC models typically predict some response variable Y as a linear function of a predictor X where the parameter $\gamma(s)$, representing the linear response of Y to changes in X, varies as a function of space:

$$Y(s) = \beta + \gamma(s)X(s) + \varepsilon(s) \quad (1)$$

where β is a global intercept, $\gamma(s)$ is the slope at spatial location s, and $\varepsilon(s)$ represents residual errors. Although $\gamma(s)$ can be approximated as a parametric function (i.e., as a function of some other covariate), a general treatment allows slope $\gamma(s)$ to vary randomly across space, $\gamma(s) = \gamma_0 + \gamma_X(s)$, where γ_0 is the

average slope across space and γ_X is a zero-mean function representing spatial variation in $\gamma(s)$. In the following, I specify that $\gamma_X(s)$ follows a zero-mean Gaussian Markov random field, $\boldsymbol{\gamma}_X \sim MVN(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\gamma}_X$ is the vector of $\gamma_X(s)$ at some set of locations and $\boldsymbol{\Sigma}$ is a spatial correlation function. This specification then allows for easy integration into existing Vector Autoregressive SpatioTemporal modeling software, available as the R package *VAST* (Thorson 2019*b*). Although I focus on implementing this approach using *VAST*, the following could also be done in any other software package that can estimate the linear interaction of a continuous variable and a spatial random field or penalized spline (e.g., package *mgcv*; Wood 2006).[1]

Interpretation of SVCs in spatiotemporal models

In the following, I will use SVCs to explain spatiotemporal variation as resulting from an annual time series. Annual time series are widely used to summarize oceanographic conditions in a given marine ecosystem, for example, where the location of the North Atlantic Oscillation is correlated with the Gulf Stream and therefore affects ocean productivity in the Northwest Atlantic shelf (O'Leary et al. 2018). I therefore describe a spatiotemporal model for a response $Y(s,t)$:

$$Y(s,t) = \beta(t) + \omega(s) + \varepsilon(s,t) \quad (2A)$$

where $\beta(t)$ is an intercept that varies for every modeled time period t (e.g., among years), $\omega(s)$ is persistent spatial variation, and $\varepsilon(s,t)$ is spatial variation that changes over time (termed "spatiotemporal" variation). Spatial shifts in $Y(s,t)$ are explained by $\varepsilon(s,t)$, such that the spatial distribution of $Y(s,t)$ could be forecasted exactly given that $\varepsilon(s,t)$ could be predicted as a known function of available covariates. In the following, I provide three alternative interpretations of SVCs.

Regression of spatiotemporal variation on an annual index

Analysts include covariates X for many different reasons, although I focus in the following on efforts to explain spatiotemporal variation $\varepsilon(s,t)$. When covariate $X(t)$ represents an annual oceanographic index, it will have the same value for all locations in a given year, such that the product $\gamma X(t)$ will explain variation in intercepts $\beta(t)$ but have no impact on spatial distribution. I therefore introduce a SVC for annual time series $X(t)$:

$$Y(s,t) = \beta(t) + \omega(s) + \varepsilon(s,t) + (\gamma_0 + \gamma_X(s))X(t) \quad (2B)$$

In many instances, intercepts $\beta(t)$ are estimated as fixed effects and in this instance, $\beta(t)$ is confounded with $\gamma_0 X(t)$ such that I specify $\gamma_0 = 0$ to allow intercepts to be identified. In other instances when $\beta(t)$ has some hierarchical structure across time (e.g., follows an autoregressive process), then $\beta(t)$ and γ_0 can be separately estimated.

Inspecting Eq. 2B, it is apparent that the product of the SVC $\gamma_0 + \gamma_X(s)$ and the annual covariate $X(t)$ generates variation $(\gamma_0 + \gamma_X(s))X(t)$ for all locations and times. Response variable $Y(s,t)$ has some true variance, which the model in Eq. 2B partitions via some combination of terms of the right side. Therefore, estimating a spatially varying impact of covariate $X(t)$ will explain some variance that would otherwise be attributed to $\varepsilon(s,t)$. In this sense, then, $(\gamma_0 + \gamma_X(s))X(t)$ defines a linear model for explaining part of the spatiotemporal residual variation in $\varepsilon(s,t)$, and including a spatially varying effect of $X(t)$ will reduce the variance of residual spatiotemporal variation $\varepsilon(s,t)$.

Random slope models

Alternatively, one can interpret SVCs as a "random slope" model. Gelman and Hill (2007) define "random intercept" models as linear models that include random variation in the intercept, that is, by interpreting $\beta + \varepsilon(s)$ in Eq. 1 as a spatially varying intercept. In contrast, "random slope" models include variation in the slope coefficient, that is, by treating $\gamma(s)$ as a spatially varying slope. Therefore, defining SVCs allows an analyst to estimate both a random-intercept and random-slope within SDMs.

Generalized analysis of nonlocal environmental impacts

This SVC model could be generalized by an approach that takes as input a spatial variable $X(s,t)$ in each year and provides as output the impact of this variable on population density at all locations. In this generalized approach, the impact for a given location s_1 depends upon the value of $X(s,t)$ everywhere, including interactions with its value at geographically distant locations, and not simply upon $X(s_1, t)$ at that single location. However, this generalized approach potentially explains complicated spatial dependencies, for example, due to the geographically distant production of phytoplankton that is then advection to a given location, and therefore requires a sufficiently flexible modeling framework. One simplification of this generalized approach is to estimate a "compression function" that maps spatial variable $X(s,t)$ to a reduced number of spatial features, and then a "projection function" that predicts the impact of $X(s,t)$ based on that reduced set of features. In this simplification, the spatially varying coefficient model can be interpreted as the projection function, while the compression function could be estimated by using an empirical orthogonal function analysis (Grimmer 1963) or other rank reduction techniques (see Appendix A in Supporting Information for more details and notation regarding this interpretation).

[1] Spatially varying coefficients can be implemented in package *mgcv* using their "formula" notation, that is, including term "s(Lat_i, Lon_i, by=Index_i)," where "Lat_i" and "Lon_i" are the latitude and longitude associated with each sample *i*, and "Index_i" is the continuous-valued variable representing the oceanographic index for each sample.

Including SVCs in *VAST*

In the following, I will analyze biomass-sampling data b_i using a delta model in the R package *VAST* (https://github.com/James-Thorson/VAST). This delta-model involves specifying the probability $p(i)$ that sample i encounters a given species, and simultaneously specifying a probability density for observed biomass given that the species is encountered:

$$\Pr(b_i = B) = \begin{cases} 1 - p(i) & \text{if } B = 0 \\ p(i) \times g\{B | r(i), \sigma_m^2\} & \text{if } B > 0 \end{cases} \quad (3)$$

where $r(i)$ is predicted biomass given an encounter, $g()$ is a probability density function for positive catches, and σ_m^2 is the residual ("measurement") variance for samples.

I specifically use a Poisson-link delta model (Thorson 2018) that simultaneously estimates two variables, numbers density N (with units numbers per unit area) and average biomass W (with units biomass per number), where the product of these two variables represents biomass density (with units biomass per unit area). The two linear predictors in the delta-model (Eq. 3) are calculated from these variables given that the assumption that individuals are randomly distributed in the vicinity of sampling:

$$p(i) = 1 - \exp(-a_i \times N(i)) \quad (4A)$$

$$r(i) = \frac{a_i \times N(i)}{p(i)} \times W(i) \quad (4B)$$

where $N(i)$ is predicted number density, a_i is the area sampled for observation i, and $W(i)$ is predicted biomass-per-individual. $N(i)$ and $W(i)$ are each specified via log-linked linear predictors:

$$\log(N(i)) = \underbrace{\beta_n(c_i, t_i)}_{\text{Intercept}} + \underbrace{\sum_{f=1}^{n_{\omega 2}} L_{\omega n}(c_i, f)\omega_n^*(s_i, f)}_{\text{Spatial variation}} + \underbrace{\sum_{f=1}^{n_{\varepsilon 2}} L_{\varepsilon n}(c_i, f)\varepsilon_n^*(s_i, f, t_i)}_{\text{Spatio-temporal variation}} + \underbrace{\sum_{p=1}^{n_p} \left(\gamma_n(c_i, t_i, p) + \sigma(c, p)\xi_n^*(s, c, p)\right) X(s_i, t_i, p)}_{\text{Density covariates}} + \underbrace{\cdots}_{\text{Other terms}} \quad (5)$$

where a similar log-linked linear predictor is defined for $\log(W(i))$. For completeness in describing the new SVC feature, I introduce notation for multivariate spatiotemporal models for each of n_c categories where *VAST* is capable of estimating a rank-reduced covariance among categories via loadings matrices $L_{\omega n}(c_i, f)$ and $L_{\varepsilon n}(c_i, f)$ for multiple "factors" f, and where $\omega_n^*(s, f)$ and $\varepsilon_n^*(s, f, t)$ are Gaussian Markov random fields representing spatial and spatiotemporal variation in $\log(N)$ at the location s_i and time t_i for each sample i. In univariate models, $L_{\omega n}(c_i, f)$ and $L_{\varepsilon n}(c_i, f)$ then become scalars representing the standard deviation of spatial or spatiotemporal variation. I configure *VAST* to use a predictive-process formulation where the value of each Gaussian Markov random field ($\xi_n^*(s, c, p)$, $\xi_w^*(s, c, p)$, $\omega_n^*(s_i, f)$, $\omega_w^*(s_i, f)$, $\varepsilon_n^*(s_i, f, t_i)$, and $\varepsilon_w^*(s_i, f, t_i)$) at each sampling location s_i is approximated as a bivariate linear interpolation from its value at three nearby "knots"; this same bilinear interpolation is used by R-INLA (Lindgren 2012), and is explained in more detail in Appendix B in Supporting Information.

VAST now includes term $\left(\gamma_n(c_i, t_i, p) + \sigma(c, p)\xi_n^*(s, c, p)\right) X(s_i, t_i, p)$ for each of n_p covariates.[2] This extended capability for density covariates includes $\xi_n^*(s, c, p)$, which is specified as a zero-mean Gaussian Markov random field with unit variance for each category c and covariate p. When estimating a SVC, *VAST* estimates its variance $\sigma^2(c, p)$ as a fixed effect and estimating $\sigma^2(c, p) = 0$ results in the reduced linear effect of that covariate. Users must specify that $\gamma_n(c_i, t_i, p) = 0$ for any covariate p that is identical across space whenever intercepts are estimated as fixed effects, and I do this when fitting to cold-pool extent; *VAST* includes user settings to specify a model where $\gamma_n(c_i, t_i, p) = 0$.

Case-study application

I explore the utility of this new SVC model using bottom-trawl survey data for fish and decapod species in the eastern Bering Sea. Previous oceanographic research has shown that the productivity and spatial structure of this marine ecosystem is strongly tied to both the spatial extent of sea ice and the seasonal timing of ice melt. Specifically, sea ice provides a spatial structure for the growth of highly productive phytoplankton, and thereby affects the timing of vertical mixing and resulting plankton community structure (Stabeno et al. 2012). Areas with sea ice are also associated with colder bottom

[2] Previous versions of *VAST* included only a conventional linear model for density covariates:

$$\log(N(i)) = \underbrace{\beta_n(c_i, t_i)}_{\text{Intercept}} + \underbrace{\sum_{f=1}^{n_{\omega 2}} L_{\omega n}(c_i, f)\omega_n(s_i, f)}_{\text{Spatial variation}} + \underbrace{\sum_{f=1}^{n_{\varepsilon 2}} L_{\varepsilon n}(c_i, f)\varepsilon_n(s_i, f, t_i)}_{\text{Spatio-temporal variation}} + \underbrace{\sum_{p=1}^{n_p} \gamma_n(c_i, t_i, p) X(s_i, t_i, p)}_{\text{Density covariates}} + \underbrace{\cdots}_{\text{Other terms}}$$

where $\gamma_n(c_i, t_i, p)$ was constant across space.

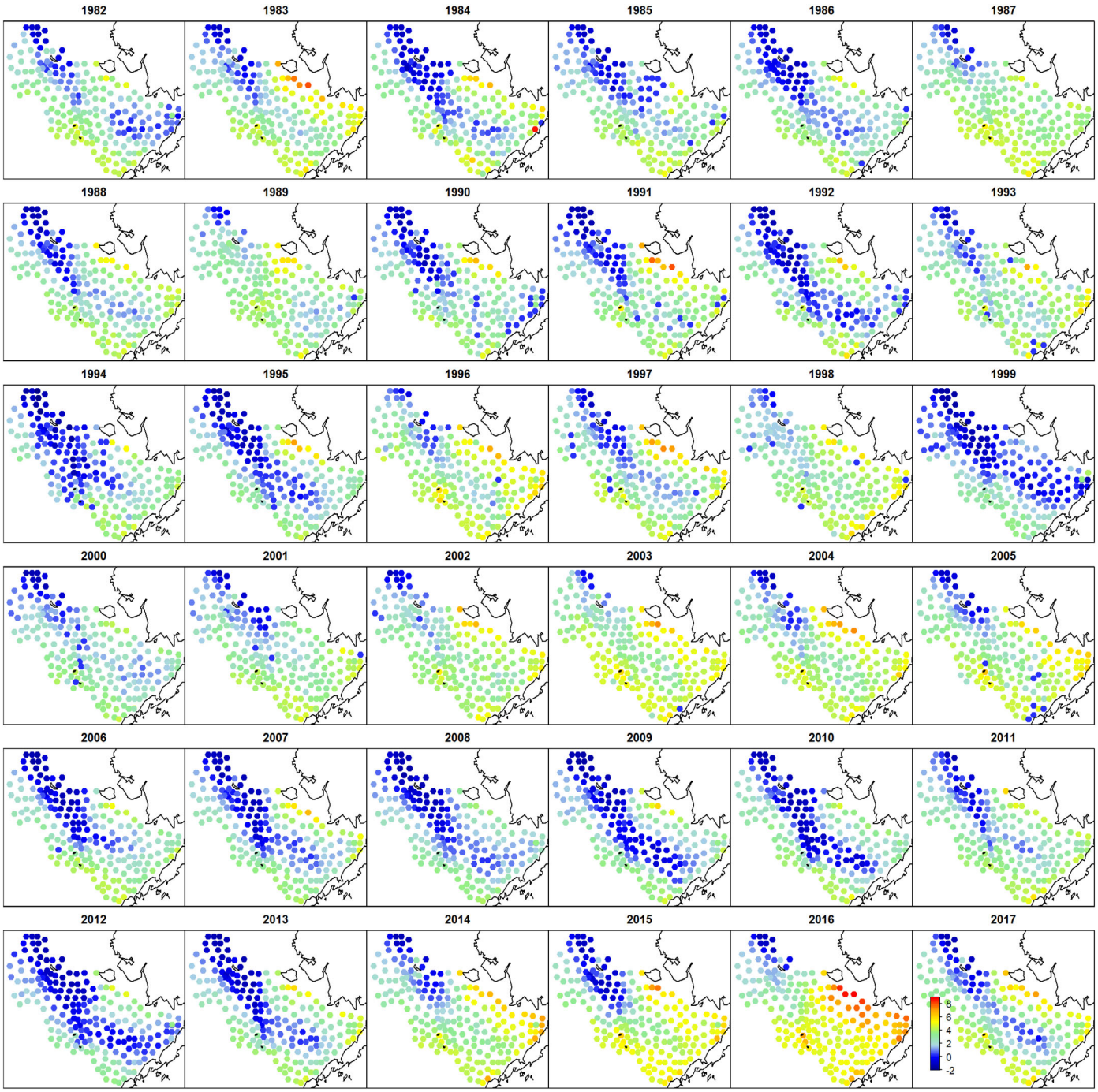

Fig. 1. Maps showing local bottom temperatures in each year 1982–2017 for each of 200 "knots" used within *VAST* to approximate spatiotemporal variation in density for each modeled species (see color bar in the bottom-right panel; blue represents low temperatures and red represents high temperatures).

temperatures, affecting distribution for marine species that either prefer or avoid near-zero water temperatures (Hunt et al. 2011). I follow past studies in designating the cold pool as the area with bottom temperatures at or below 2°C (Fig. 1 for the annual landscape of summertime bottom temperatures), and explore using the spatial extent of this cold pool as an annual oceanographic index to describe spatiotemporal variation in distribution. The index of cold-pool extent (Fig. 2) is based on net sensor measurements in the eastern Bering Sea shelf survey (Lauth and Conner 2016) and has been published previously (e.g., Thorson et al. 2017) but is here updated to include data for 2016–2017 (Bob Lauth, pers. comm.).

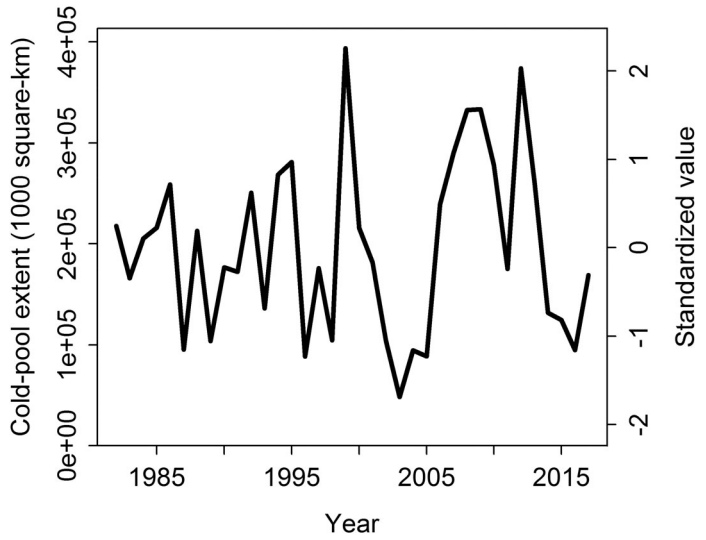

Fig. 2. Time series showing extent of cold pool in the eastern Bering Sea for modeled years, 1982–2017 (*x*-axis), where the cold-pool extent is measured as the area with bottom temperatures at or below 2 °C. The time series is labeled both for the measured value (left *y*-axis label) and the standardized value (right *y*-axis label), where the standardized value is included in the model.

Specifically, I seek to answer: (1) does a spatially varying coefficient model for cold-pool extent explain a substantial portion of variation in spatial distribution for species in this region? And (2) does a spatially varying impact of the cold pool remain useful and significant even when also including local temperature as a covariate? To answer these questions, I use data for the 17 numerically dominant bottom-associated fish and decapod species in the eastern Bering Sea slope bottom trawl survey conducted annually from 1982 onward using standardized bottom trawl gear by the Alaska Fisheries Science Center (Lauth and Conner 2016). This survey uses a fixed station design involving nearly 370 stations on a 20 km by 20 km grid (including some areas with more dense sampling near significant islands), and data are publicly available online (http://www.afsc.noaa.gov/RACE/groundfish/survey_data/data.htm). I explore the potential effect of local environmental conditions (a quadratic effect of bottom temperature) and regional conditions (a spatially varying effect of cold-pool extent), and this results in the following reduced model:

$$\log(N(i)) = \underbrace{\beta_n(t_i)}_{\text{Intercept}} + \underbrace{\sigma_{\omega n}\omega_n^*(s_i)}_{\text{Spatial variation}} + \underbrace{\sigma_{\varepsilon n}\varepsilon_n^*(s_i,t_i)}_{\text{Spatio-temporal variation}} + \underbrace{\gamma_n(t_i,1)T(s_i,t_i) + \gamma_n(t_i,2)T^2(s_i,t_i)}_{\text{Quadratic effect of bottom temperature}} + \underbrace{\sigma_{\xi n}\xi_n^*(s,p)C(t_i)}_{\text{Effect of cold pool}} \quad (6)$$

where $T(s,t)$ is local bottom temperature for each location s_i and year t_i for sample i, $C(t)$ is cold-pool extent for year t, and a similar model is again specified for average density, $\log(W(i))$.

For each species, I fit eight models, formed as a 4 by 2 factorial cross of four model structures and two data sets. The four model structures include:

1. *No covariates*: A model with spatial and spatiotemporal variation for both numbers density $N(i)$ and average weight $W(i)$ but no habitat covariates (i.e., $\gamma_n = \gamma_w = 0$ and $\sigma_{\xi n} = \sigma_{\xi w} = 0$).
2. *Local temperature*: Identical to *No covariates* except also including a quadratic effect of bottom temperature, representing a dome-shaped response of local density to local temperatures (i.e., fixing $\sigma_{\xi n} = \sigma_{\xi w} = 0$).
3. *Cold-pool extent*: Identical to *No covariates* except also including a SVC linking numbers density $N(i)$ and average weight $W(i)$ to cold-pool extent $C(t)$ (i.e., fixing $\gamma_n = \gamma_w = 0$).
4. *Both*: Identical to *No covariates* except including both a quadratic effect of local temperature and a spatially varying effect of cold-pool extent (i.e., the full model in Eq. 6).

For all models, I specify a first-order autoregressive process for spatiotemporal variation, such that hotspots in density are predicted to persist for subsequent years, where the degree of autoregression is estimated from available data. I then calculate the proportion of residual spatial and spatiotemporal variation in numbers density and average weight that is explained by local temperature, cold-pool extent, or both simultaneously (Appendix C in Supporting Information) and use this to determine whether the model with both temperature and cold-pool extent has increased explanatory power relative to the model with just local temperature effects.

I fit each of these four models to two different data sets for each species:

A. *Full data*: I fit to all bottom-trawl data 1982–2017 available online in January 2018. Given that I am predicting density for this same set of years 1982–2017, I do not need to impose any restrictions of model intercepts, $\beta_n(t)$ and $\beta_w(t)$, and treat them as fixed effects.
B. *Reduced data*: I fit to data 1982–2014 and forecast future distribution shifts 2015–2017. I conduct this forecast under the assumption of perfect information regarding the value for cold-pool extent (i.e., using subsequent observations as forecasted values of ocean conditions in those future years); the use of perfect information for covariate $X(t)$ implies that results represent a best-case scenario for the performance of this forecast (i.e., a real-world application would include additional error in $X(t)$ during forecast years which I do not explore here). Given that I have no data to inform intercept estimates for forecast years, I assume that model intercepts follow a probability distribution, which I specify as following a random-walk process where the variance of this random walk is estimated from available data.

I fit all models using release 2.2.0 of the R package *VAST* (Thorson and Barnett 2017; Thorson 2019*b*) using Microsoft Open R (R Core Team 2017). *VAST* estimates fixed effects

using *TMB* (Kristensen et al. 2016) and also applies a stochastic partial differential equation approximation to a Matern spatial correlation function (Lindgren et al. 2011) to reduce computational requirements while also using a predictive-process formulation to interpolation spatial variation between a pre-specified number of "knots" (see Appendix B in Supporting Information for more details). I use 200 knots while confirming that results are qualitatively similar even when changing this number, and choose the location of knots using a k-means algorithm that minimizes the average distance between each sample and the nearest knot. I calculate standard errors using a Bayesian generalization of the delta-method (Kass and Steffey 1989).

Results

Plotting the spatially varying impact of cold-pool extent (ξ_n and ξ_w) shows how distribution is expected to differ in years with high cold-pool extent relative to years with average conditions. Visualizing this effect for four biologically diverse species shows the varied region-scale impacts of cold-pool extent on species distribution in the eastern Bering Sea (Fig. 3). *G. chalcogrammus* shows decreased densities in northern inner domain during years with high cold-pool extent, and cold-pool extent has a larger average effect on average weight W than numbers density N (i.e., the scale of variation for $\xi_w(s)$ in the top-right panel of Fig. 3 is larger than for $\xi_n(s)$ in the top-left). *Hippoglossus stenolepsis* shows elevated number density (and therefore encounter probability) in the middle domain during years with high cold-pool extent, but also elevated average biomass offshore from Bristol Bay. By contrast, *Atheresthes stomias* shows elevated number density and average biomass in the outer domain during years with large cold-pool extent, and *Chionoecetes bairdi* shows relatively little response in number density (and therefore encounter probability) in years with a large cold pool (i.e., the color scale in the bottom-left panel of Fig. 3 is very small), while average biomass is somewhat elevated in small concentrations in the southern middle domain during these years. These responses are just a few of the varied impacts occurring across all 17 species analyzed here (Appendix D in Supporting Information).

Cold-pool extent explains a portion of variation in log-density for species in the eastern Bering Sea even when including a quadratic effect of bottom temperature (Fig. 4), although the standard deviation of variation explained by bottom temperature is nearly twice as great on average (0.26/0.21; Fig. 4, second row) as the spatially varying impact of cold-pool extent (0.10/0.11; Fig. 4, first row) when both effects are estimated simultaneously. Including both bottom temperature and cold-pool extent explains 9% of the residual variation over time for numbers density (Fig. 4, bottom-left panel) and 14% of residual variation over time for average biomass (Fig. 4, bottom-right panel). Local temperature in isolation explains only 6% and 8% of residual variation in number density and average biomass, respectively, and both cold-pool extent and bottom temperature contribute substantially to the overall reduction in variance (i.e., as shown by the similar reduction for "Temp" and "ColdPool" models in bottom row of Fig. 4).

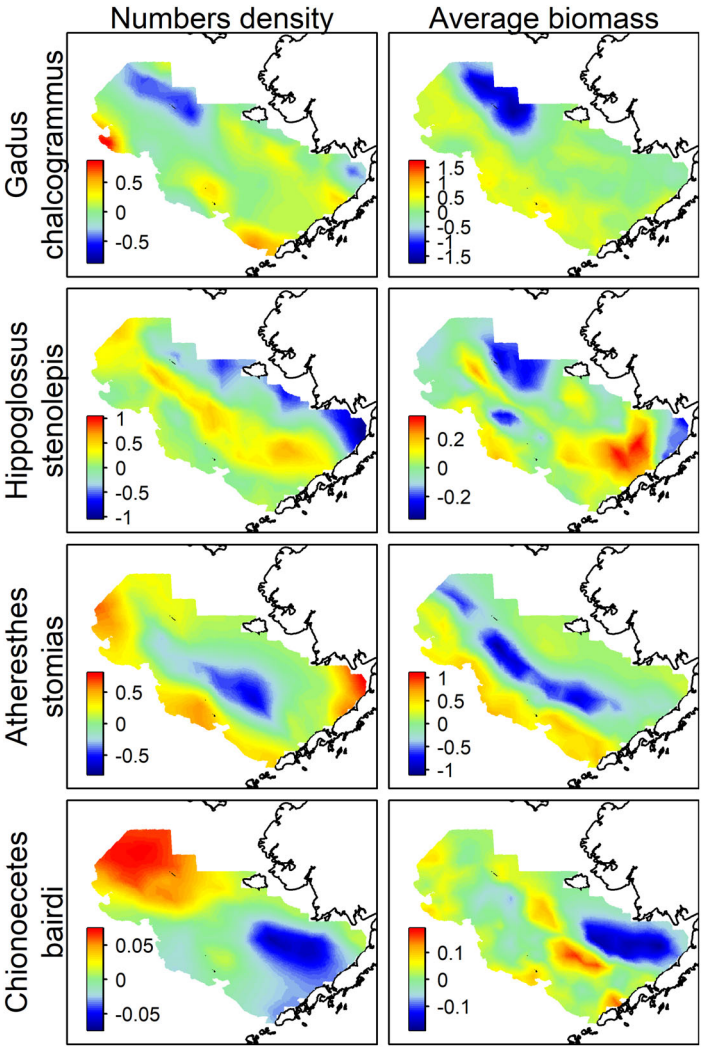

Fig. 3. Illustration of the spatially varying coefficients $\xi_n(s)$ and $\xi_w(s)$ that, respectively, link annual cold-pool extent to either numbers density N (left column) or average biomass W (right column) in the Poisson-link delta model used to predict biomass for each species in the eastern Bering Sea (see Eqs. 3–5 for details). Although variables N and W have units numbers-per-area and biomass-per-number, respectively, the spatially varying coefficient affects these via a log-link and therefore is dimensionless. Results are shown for the model including a spatially varying impact of cold-pool extent but not a quadratic effect of local bottom temperature. Cold-pool extent is standardized to have a mean of zero and standard deviation of one prior to use and all results are using a Poisson-link delta model such that cold-pool extent has a linear effect on log-biomass density; consequently, a location with a coefficient of 0.1 indicates an approximately 10% increase in expected density for every 1 standard deviation (87,504 km^2) increase in cold-pool extent. Note that the color scale differs for each species (rows, labeled on left) and linear predictor (see color bar for scale in each panel).

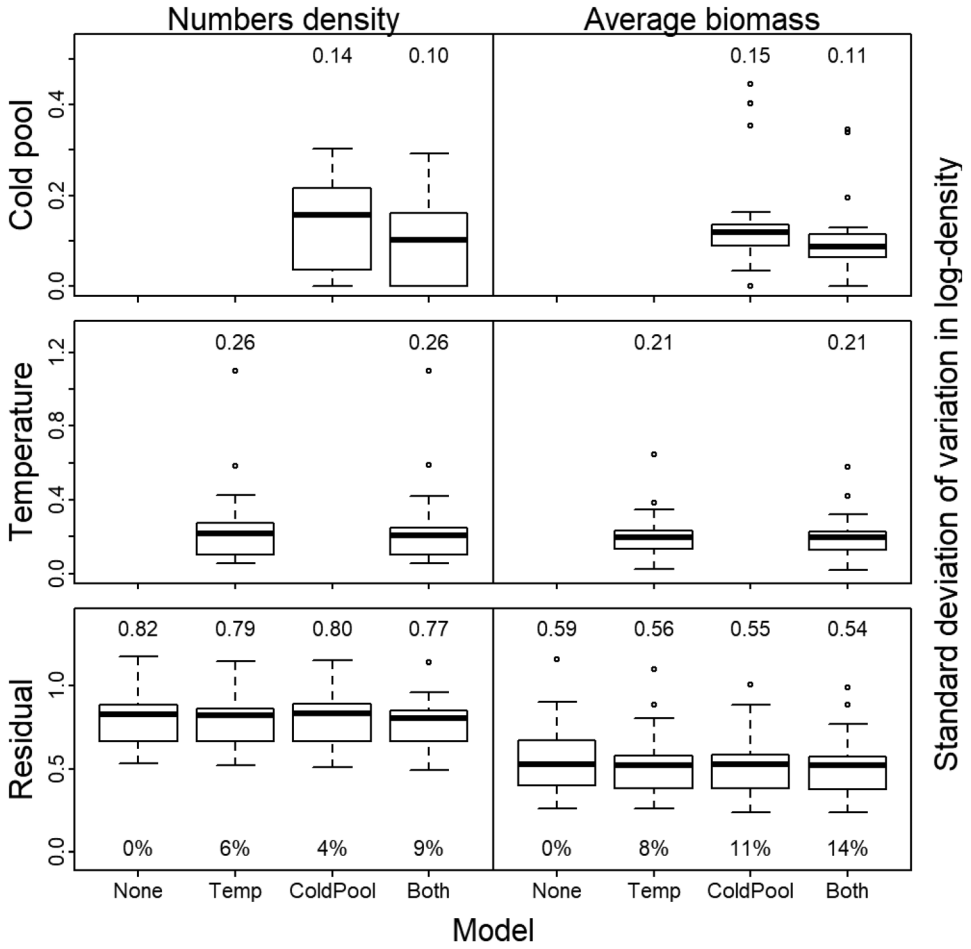

Fig. 4. Summary of the standard deviation of variation in log-density caused by a spatially varying link to cold-pool extent (first row), a quadratic response to bottom temperature (second row), or residual variation for site-specific density over time (third row) as impacting either numbers density (left column) or average biomass (right column); see Appendix C in Supporting Information for detailed calculation. Each panel shows a boxplot (line: median; box: interquartile range; whiskers: the furthest point within 1.5 times the interquartile range from the median) summarizing the estimated standard deviation across all 17 analyzed species for a given model (None: neither temperature nor cold pool; Both: both temperature and cold-pool responses). All panels list the average standard deviation across all 17 species above each boxplot, and the bottom row also lists the average reduction in residual variance relative to the model without covariates (where, e.g., 50% would indicate that covariates explain half of spatiotemporal variance). Note the different *y*-axis scale for each row.

Relatedly, the Akaike Information Criterion (AIC) indicates that including both local bottom temperature and annual cold-pool extent is parsimonious for 13 of 17 species (Table 1), and selects local temperature or cold-pool extent for all species except *Hyas lyratus*. Comparing a model with cold-pool extent to the AIC selected model (i.e., comparing "Both" with "Temp" for *Chionoecetes opilio* or *Lycodes brevipes*) shows that including the cold-pool degrades model parsimony very little even when it is not selected, and exploration shows that the estimate of $\sigma_{\xi 1}$ and $\sigma_{\xi 2}$ are approaching zero in these instances.

Finally, examining forecasts of northward center of gravity when fitting to biomass-sampling data for 1982–2014 and forecasting distribution shifts in 2015–2017 (Fig. 5) shows instances where incorporating local temperature and cold-pool extent improves forecasts of distribution shift from 2015 to 2017. For example, *Lycodes palearis* shows a southward shift in 2015 followed by a rapid shift 100 km northward in 2016–2017. This northward shift is forecasted by all models because the species has a more southward distribution in 2014 than its long-term average and the autocorrelation in spatiotemporal variation in all models causes it to revert to its long-term average. However, models including local temperature forecast a substantial northward shift in 2015–2016, in accordance with subsequent measurements. Local temperature is similarly helpful in forecasting the northward shift for *H. stenolepsis* and *Hyas coarctatus* in 2016. Meanwhile, cold-pool extent is helpful in improving forecasts of northward distribution in 2015 for *Gadus macrocephalus* relative to models with only local temperature effects. Comparing the error in

Table 1. Comparison of ΔAIC for 17 species (listed by common and scientific name) when biomass-sampling data are analyzed using four candidate models (the model with lowest AIC is indicated in bold); Temp: a quadratic effect of local bottom temperature on number density and average biomass; ColdPool: a spatially varying effect of annual cold-pool extent on number density and average biomass; None: neither temperature nor cold-pool effects; Both: both temperature and cold-pool effects.

Scientific name	Common name	None	Temp	ColdPool	Both
Gadus chalcogrammus	Walleye pollock	239.4	68.2	138.5	**0.0**
Gadus macrocephalus	Pacific cod	528.6	134.4	363.8	**0.0**
Hippoglossoides elassodon	Flathead sole	175.5	6.2	142.6	**0.0**
Chionoecetes opilio	Snow crab	38.4	**0.0**	37.8	0.5
Hippoglossus stenolepis	Pacific halibut	260.9	87.5	178.4	**0.0**
Limanda aspera	Yellowfin sole	79.8	6.4	67.4	**0.0**
Pleuronectes quadrituberculatus	Alaska plaice	70.3	37.4	20.7	**0.0**
Chionoecetes bairdi	Tanner crab	0.8	6.5	0.0	5.8
Podothecus accipenserinus	Sturgeon poacher	212.3	7.4	157.5	**0.0**
Atheresthes stomias	Arrowtooth flounder	475.4	34.6	365.8	**0.0**
Hyas coarctatus		31.5	11.0	17.3	**0.0**
Myoxocephalus polyacanthocephalus	Great sculpin	85.5	19.5	47.3	**0.0**
Lycodes palearis	Wattled eelpout	98.4	8.5	62.4	**0.0**
Myoxocephalus jaok	Plain sculpin	104.5	26.4	70.1	**0.0**
Hyas lyratus		**0.0**	4.3	2.2	6.7
Paralithodes camtschaticus	Red king crab	32.4	9.1	23.1	**0.0**
Lycodes brevipes	Shortfin eelpout	8.1	**0.0**	11.2	2.8

forecasted center of gravity among models (Fig. 6) shows that including temperature and cold-pool extent decreases the median absolute error by 33% (from 15.3 to 10.2 km) for 1-yr forecasts, 39% for 2-yr forecasts (from 32.3 to 19.8 km), and 6% for 3-yr forecasts (from 38.1 to 36.0 km) relative to a model with neither covariate, and similarly decreases bias in forecasts for each year. The model without covariates itself has similar or lower errors than the persistence forecast for 2- and 3-yr forecasts, due to its inclusion of density dependence as represented by the autoregressive structure on density hotspots.

Discussion

In this study, I have shown that SVC models can be used to explain changes in spatial distribution using annual oceanographic indices. Using data for 17 bottom-associated fish and crab species in the eastern Bering Sea, I have shown that a dome-shaped response to local bottom temperature explains 6%–8% of spatiotemporal variation in two predictors of density, while local temperature and the additional spatially varying impact of cold-pool extent explain 9%–14% of variation in these two predictors. In addition, both local temperature and cold-pool effects are identified as parsimonious by AIC for 13 of the 17 species, and including both effects improves 1-, 2-, or 3-yr forecasts of recent distribution shifts over a model without covariates or a persistence forecast. I therefore conclude that SVCs are a useful tool in the toolbox of methods available for explaining and forecasting distribution within SDMs, and are particularly valuable as a way to integrate regional oceanographic indicators. However, I note that forecast results presented here represent an "upper bound" on forecast skill, because real-world applications must also forecast oceanographic variables (e.g., cold-pool extent) and biological responses simultaneously, whereas I have instead used subsequent observations as perfect information for the oceanographic variable. Improving skill for forecasts of cold-pool extent in the eastern Bering Sea is a topic of active research (e.g., Hermann et al. 2016), but current research is focused on 9 month forecasts rather than the 2- and 3-yr forecasts explored here. The decision of how far to forecast future distribution shifts in other systems will depend upon available skill in forecasting oceanographic variables as well as the extent of skill gained by modeling short-term density dependence in biological dynamics (e.g., Thorson 2019*a*).

Given the benefit of including spatially varying effects of oceanographic indices, I also highlight two paradoxes of the results. Specifically, cold-pool extent is calculated from local measurements of bottom temperature and represents a highly compressed representation of physical conditions which "loses" information about annual differences in the spatial distribution of cold habitats. However, the analysis shows that the spatially varying impact of cold-pool extent contains additional information about species distribution beyond the information contained in local temperatures. How do fishes "know" about temperatures occurring in habitats outside the

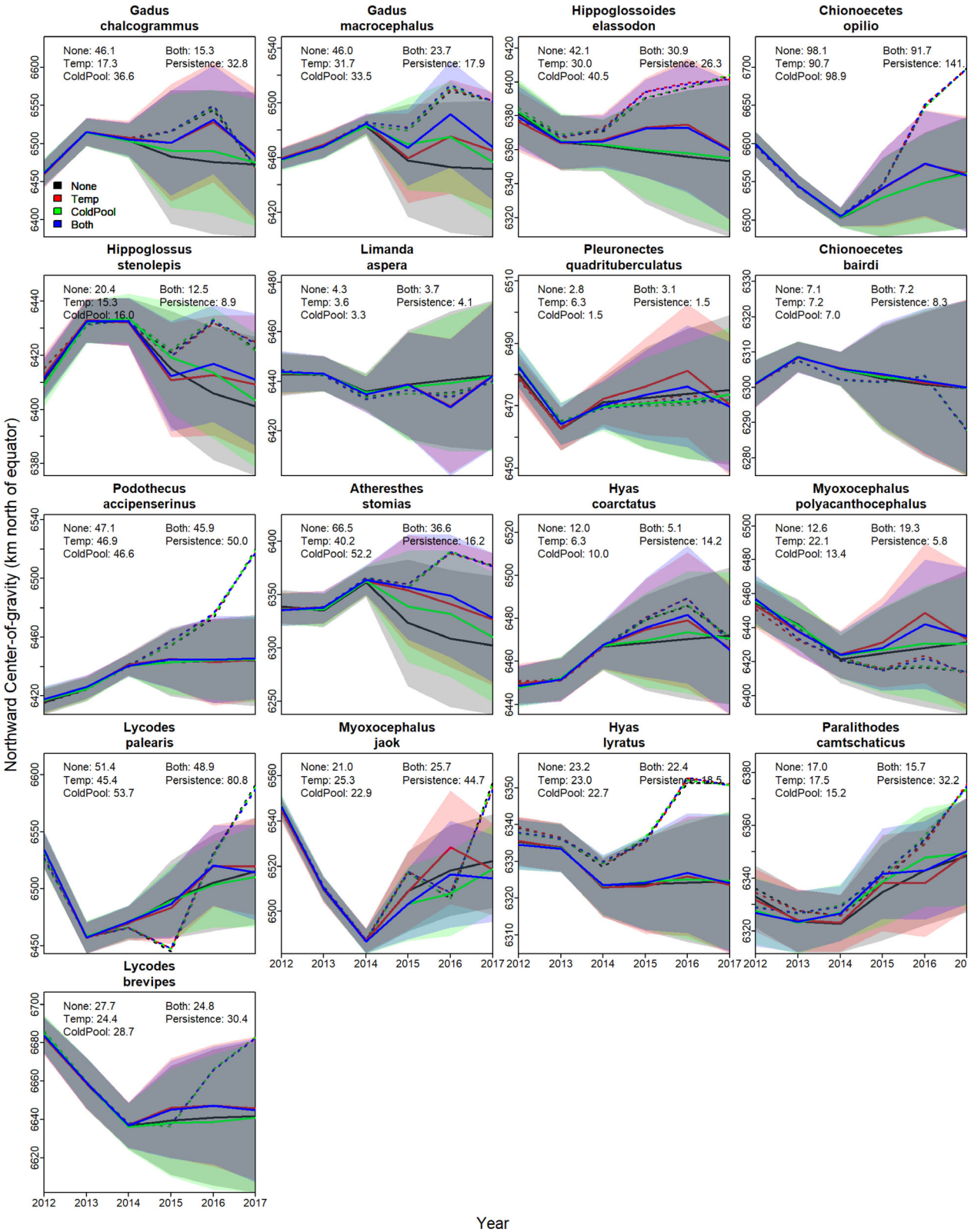

Fig. 5. Estimates of northward center of gravity (*y*-axis; note different scale for each species) for each year 2012–2017 (*x*-axis) for each of 17 species (panels) and four models (line: estimate; shaded area: +/− one standard error; see color labels in top-left panel) when fitting to data 1982–2014 and forecasting distribution in 2015–2017. The dashed lines show estimates when fitting to all data (2012–2017) for each model, which are generally very similar. Each panel includes the average error for years 2015–2017 (computed as the difference between forecasts and estimates using all data when averaging across models) for each of four models as well as a "persistence" forecast (i.e., forecasting that the distribution does not change after 2014).

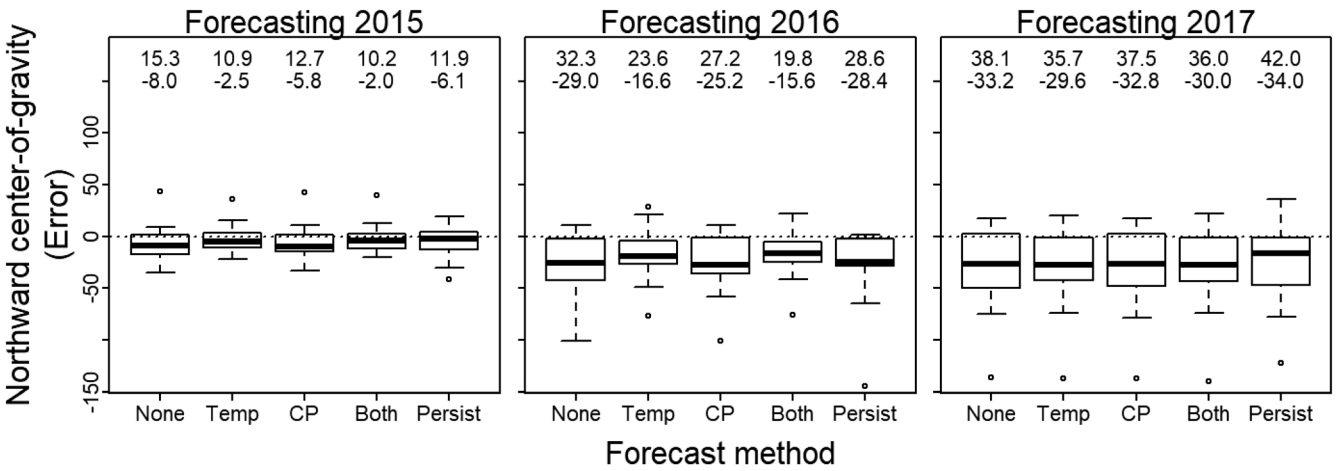

Fig. 6. Distribution of error for forecasts of northward center of gravity (computed as the difference between forecasts and estimates using all data when averaging across models) for all 17 species when fitting to data 1982–2014 and forecasting distribution in 2015 (left panel), 2016 (middle panel), or 2017 (right panel) using four models or a "persistence" forecast (see Fig. 5 caption for details). The top of each panel lists the median absolute error (a value close to zero is better) and bias (a value close to zero is better, where a negative value indicates an estimate that is more southward than subsequent observations) for each forecast model.

area being sampled? And how does compressing temperature measurements result in "more information?" I address these two questions below.

Regarding how fishes "know" about temperatures occurring in habitats outside the area being sampled: the fish and decapods analyzed here are highly mobile and select their habitats after having "sampled" (migrated through or foraged within) other habitats available regionally. This answer is obvious biologically, and community ecologists acknowledge that many nonlocal and regional mechanisms are needed to explain species distribution and density (HilleRisLambers et al. 2013; Heino et al. 2017). However, SDMs have typically focused on including local habitat measurements within regression models for predicting local densities, and these nonlocal mechanisms are rarely included explicitly in SDMs. Alternatively, some SDMs apply a kernel smoother to local covariates prior to inclusion (Chandler and Hepinstall-Cymerman 2016; Frishkoff et al. 2017); this smoother captures the response to a spatial average of nearby habitats, but does not allow for complicated dependencies on spatially distant habitats like the SVC model explored here. I therefore argue that SVCs for annual oceanographic indices represent a useful and flexible way to represent habitat selection as animals respond to information about resource availability in years with different oceanographic conditions. For example, pollock in the Bering Sea have been shown to distribute along the sea-ice break during spring (De Robertis and Cokelet 2012), and the location of the ice-break is driven by several region-scale and lagged processes that are not necessarily related to local bottom temperatures (Stabeno et al. 2007). In these cases, a regional covariate (e.g., sea-ice extent) represents the integrated effect of lagged and geographically distant habitat conditions that affect fish decisions regarding habitat selection.

Regarding how compressing temperature measurements result in more information: the spatially varying effect of cold-pool extent upon number density and average weight (Fig. 3) differs greatly among species in a way that cannot be explained by the local impact of bottom temperature. For example, *G. chalcogrammus* has reduced densities along the northern boundary of the eastern Bering Sea in years with high cold-pool extent, and this clearly mirrors the areas where bottom temperatures are strongly impacted by cold-pool extent. However, *Clupea pallasii* and *H. stenolepsis* show contrasting effects of cold-pool extent and neither exactly coincides with areas that show greatest changes in local temperature. I therefore see a spatially varying response to cold-pool extent as an approximation to regional temperature impacts that vary among species. In particular, compressing temperature to calculate cold-pool extent and then using a species-specific projection of this index on distribution approximates a flexible process for translating both local and nonlocal temperatures to densities maps (see Appendix A in Supporting Information for more details). I therefore encourage ongoing research to allow spatiotemporal models to simultaneously estimate a generalized function calculating regional indices from local environmental conditions, in addition to the nonparametric function that SVC models already involve. This could perhaps be accomplished by simultaneously estimating parameters for the empirical orthogonal function models that are typically used to calculate oceanographic indices and the SDMs that is used to estimate species distributions.

In particular, a spatially varying coefficient for regional oceanographic conditions allows analysts to assimilate both fine-scale and regional covariates in a single modeling framework, rather than building separate models for each resolution

(e.g., Smart et al. 2012). Alternatively, previous studies have included regional covariates by estimating a separate spatial distribution for "warm" and "cool" years in the eastern Bering Sea (e.g., Hollowed et al. 2012), while the SVC model avoids classifying regional conditions into categories prior to analysis. In the eastern Bering Sea, cold-pool extent is associated with many processes that might be associated mechanistically with habitat selection and resulting fish distribution including: plankton prey densities (Stabeno et al. 2012), dominant modes of physical transport (Zhang et al. 2012), and partitioning of available primary productivity between pelagic and demersal communities (Hunt et al. 2011). The SVC allows these many upstream and downstream processes affecting distribution to be summarized in an annual index that is used to hindcast and forecast spatial patterns in community structure.

Most importantly, I see SVCs for oceanographic indices as an important programmatic tool to translate climatological research from physical oceanography into the distribution models that are used for marine spatial planning (Rassweiler et al. 2014), essential fish habitat (Rooney et al. 2018), and EBM (Link and Browman 2014). Methods exist for translating indices like the PDO into stock assessment models used for defining harvest limits (Schirripa et al. 2009), and these methods are increasingly used to evaluate ecological claims regarding bottom-up limits for fisheries productivity (O'Leary et al. 2018). However, oceanographic indices are less commonly used in SDM forecasts of distribution shift. I hope that demonstrating the benefit of oceanographic indices in SDMs will continue to "build a bridge" between fisheries oceanography and EBM.

References

Araújo, M. B., and M. New. 2007. Ensemble forecasting of species distributions. Trends Ecol. Evol. **22**: 42–47. doi:10.1016/j.tree.2006.09.010

Bacheler, N. M., K. M. Bailey, L. Ciannelli, V. Bartolino, and K.-S. Chan. 2009. Density-dependent, landscape, and climate effects on spawning distribution of walleye pollock *Theragra chalcogramma*. Mar. Ecol. Prog. Ser. **391**: 1–12. doi:10.3354/meps08259

Bacheler, N. M., L. Ciannelli, K. M. Bailey, and V. Bartolino. 2012. Do walleye pollock exhibit flexibility in where or when they spawn based on variability in water temperature? Deep Sea Res. Part II Top. Stud. Oceanogr. **65–70**: 208–216. doi:10.1016/j.dsr2.2012.02.001

Barbeaux, S. J., and A. B. Hollowed. 2018. Ontogeny matters: Climate variability and effects on fish distribution in the eastern Bering Sea. Fish. Oceanogr. **27**: 1–15. doi:10.1111/fog.12229

Bartolino, V., L. Ciannelli, N. M. Bacheler, and K.-S. Chan. 2011. Ontogenetic and sex-specific differences in density-dependent habitat selection of a marine fish population. Ecology **92**: 189–200. doi:10.1890/09-1129.1

Chandler, R., and J. Hepinstall-Cymerman. 2016. Estimating the spatial scales of landscape effects on abundance. Landsc. Ecol. **31**: 1383–1394. doi:10.1007/s10980-016-0380-z

Cheung, W. W. L., V. W. Y. Lam, and D. Pauly. 2008. Modelling present and climate-shifted distribution of marine fishes and invertebrates. Fisheries Centre, Univ. of British Columbia. doi:10.1021/la802452v

De Robertis, A., and E. D. Cokelet. 2012. Distribution of fish and macrozooplankton in ice-covered and open-water areas of the eastern Bering Sea. Deep Sea Res. Part II Top. Stud. Oceanogr. **65–70**: 217–229. doi:10.1016/j.dsr2.2012.02.005

Dolder, P. J., J. T. Thorson, and C. Minto. 2018. Spatial separation of catches in highly mixed fisheries. Sci. Rep. **8**: 13886. doi:10.1038/s41598-018-31881-w

Finley, A. O. 2011. Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence. Methods Ecol. Evol. **2**: 143–154. doi:10.1111/j.2041-210X.2010.00060.x

Frishkoff, L. O., D. L. Mahler, and M.-J. Fortin. 2017. Integrating over uncertainty in spatial scale of response within multispecies occupancy models yields more accurate assessments of community composition. bioRxiv. doi:10.1101/143669

Fulton, E. A., and others. 2011. Lessons in modelling and management of marine ecosystems: The Atlantis experience. Fish Fish. **12**: 171–188. doi:10.1111/j.1467-2979.2011.00412.x

Gelfand, A. E., A. M. Schmidt, S. Banerjee, and C. F. Sirmans. 2004. Nonstationary multivariate process modeling through spatially varying coregionalization. Test **13**: 263–312. doi:10.1007/BF02595775

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge Univ. Press. doi:10.1016/S0072-9752(07)85018-4

Grimmer, M. 1963. The space-filtering of monthly surface temperature anomaly data in terms of pattern, using empirical orthogonal functions. Q. J. Roy. Meteorol. Soc. **89**: 395–408. doi:10.1002/qj.49708938111

Heino, J., J. Soininen, J. Alahuhta, J. Lappalainen, and R. Virtanen. 2017. Metacommunity ecology meets biogeography: Effects of geographical region, spatial dynamics and environmental filtering on community structure in aquatic organisms. Oecologia **183**: 121–137. doi:10.1007/s00442-016-3750-y

Hermann, A. J., and others. 2016. Projected future biophysical states of the Bering Sea. Deep Sea Res. Part II Top. Stud. Oceanogr. **134**: 30–47. doi:10.1016/j.dsr2.2015.11.001

HilleRisLambers, J., M. A. Harsch, A. K. Ettinger, K. R. Ford, and E. J. Theobald. 2013. How will biotic interactions influence climate change–induced range shifts? Ann. N. Y. Acad. Sci. **1297**: 112–125. doi:10.1111/nyas.12182

Hollowed, A. B., S. J. Barbeaux, E. D. Cokelet, E. Farley, S. Kotwicki, P. H. Ressler, C. Spital, and C. D. Wilson. 2012.

Effects of climate variations on pelagic ocean habitats and their role in structuring forage fish distributions in the Bering Sea. Deep Sea Res. Part II Top. Stud. Oceanogr. **65–70**: 230–250. doi:10.1016/j.dsr2.2012.02.008

Hunt, G. L., and others. 2011. Climate impacts on eastern Bering Sea foodwebs: A synthesis of new data and an assessment of the oscillating control hypothesis. ICES J. Mar. Sci. **68**: 1230–1243. doi:10.1093/icesjms/fsr036

Kai, M., J. T. Thorson, K. R. Piner, and M. N. Maunder. 2017. Spatio-temporal variation in size-structured populations using fishery data: An application to shortfin mako (*Isurus oxyrinchus*) in the Pacific Ocean. Can. J. Fish. Aquat. Sci. **74**: 1765–1780. doi:10.1139/cjfas-2016-0327

Karp, M. A., et al. 2019. Accounting for shifting distributions and changing productivity in the development of scientific advice for fishery management. ICES J. Mar. Sci. doi:10.1093/icesjms/fsz048

Kass, R. E., and D. Steffey. 1989. Approximate bayesian inference in conditionally independent hierarchical models (parametric empirical bayes models). J. Am. Stat. Assoc. **84**: 717–726. doi:10.2307/2289653

Kidson, J. W. 1975. Tropical eigenvector analysis and the southern oscillation. Mon. Weather Rev. **103**: 187–196. doi:10.1175/1520-0493(1975)103<0187:TEAATS>2.0.CO;2

Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. M. Bell. 2016. TMB: Automatic differentiation and laplace approximation. J. Stat. Softw. **70**: 1–21. doi:10.18637/jss.v070.i05

Kristensen, K., U. H. Thygesen, K. H. Andersen, and J. E. Beyer. 2014. Estimating spatio-temporal dynamics of size-structured populations. Can. J. Fish. Aquat. Sci. **71**: 326–336. doi:10.1139/cjfas-2013-0151

Lauth, R. R., and J. Conner. 2016. Results of the 2013 eastern Bering Sea continental shelf bottom trawl survey of groundfish and invertebrate resources. NOAA Technical Memorandum NMFS-AFSC-331. NMFS-AFSC-331. doi:10.1021/acs.jpclett.6b01835

Lindgren, F. 2012. Continuous domain spatial models in R-INLA. ISBA Bull. **19**: 14–20.

Lindgren, F., H. Rue, and J. Lindström. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B Stat. Methodol. **73**: 423–498. doi:10.1111/j.1467-9868.2011.00777.x

Link, J. S., and H. I. Browman. 2014. Integrating what? Levels of marine ecosystem-based assessment and management. ICES J. Mar. Sci. **71**: 1170–1173. doi:10.1093/icesjms/fsu026

Mantua, N. J., and S. R. Hare. 2002. The Pacific decadal oscillation. J. Oceanogr. **58**: 35–44. doi:10.1023/A:1015820616384

North Pacific Fishery Management Council. 2019. Bering Sea Fishery Ecosystem Plan. North Pacific Fishery Management Council. doi:10.3352/jeehp.2019.16.16

O'Leary, C. A., T. J. Miller, J. T. Thorson, and J. A. Nye. 2018. Understanding historical summer flounder (*Paralichthys dentatus*) abundance patterns through the incorporation of oceanography-dependent vital rates in Bayesian hierarchical models. Can. J. Fish. Aquat. Sci. doi:10.1139/cjfas-2018-0092

Perretti, C. T., and J. T. Thorson. 2019. Spatio-temporal dynamics of summer flounder (*Paralichthys dentatus*) on the Northeast US shelf. Fish. Res. **215**: 62–68. doi:10.1016/j.fishres.2019.03.006

Pinsky, M. L., G. Reygondeau, R. Caddell, J. Palacios-Abrantes, J. Spijkers, and W. W. L. Cheung. 2018. Preparing ocean governance for species on the move. Science **360**: 1189–1191. doi:10.1126/science.aat2360

R Core Team. 2017. R: A language and environment for statistical computing, R Foundation for Statistical Computing

Rassweiler, A., C. Costello, R. Hilborn, and D. A. Siegel. 2014. Integrating scientific guidance into marine spatial planning. Proc. R. Soc. B Biol. Sci. **281**: 20132252.

Rooney, S. C., C. N. Rooper, E. A. Laman, K. A. Turner, D. W. Cooper, and M. Zimmermann. 2018. Model-based essential fish habitat definitions for Gulf of Alaska groundfish species. NOAA Fisheries. doi:10.1097/ICB.0000000000000841

Rooper, C. N., M. F. Sigler, P. Goddard, P. Malecha, R. Towler, K. Williams, R. Wilborn, and M. Zimmermann. 2016. Validation and improvement of species distribution models for structure-forming invertebrates in the eastern Bering Sea with an independent survey. Mar. Ecol. Prog. Ser. **551**: 117–130. doi:10.3354/meps11703

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

Smart, T. I., J. T. Duffy-Anderson, J. K. Horne, E. V. Farley, C. D. Wilson, and J. M. Napp. 2012. Influence of environment on walleye pollock eggs, larvae, and juveniles in the southeastern Bering Sea. Deep Sea Res. Part II Top. Stud. Oceanogr. **65–70**: 196–207. doi:10.1016/j.dsr2.2012.02.018

Stabeno, P. J., N. A. Bond, and S. A. Salo. 2007. On the recent warming of the southeastern Bering Sea shelf. Deep Sea Res. Part II Top. Stud. Oceanogr. **54**: 2599–2618. doi:10.1016/j.dsr2.2007.08.023

Stabeno, P. J., N. B. Kachel, S. E. Moore, J. M. Napp, M. Sigler, A. Yamaguchi, and A. N. Zerbini. 2012. Comparison of warm and cold years on the southeastern Bering Sea shelf and some implications for the ecosystem. Deep Sea Res. Part II Top. Stud. Oceanogr. **65–70**: 31–45. doi:10.1016/j.dsr2.2012.02.020

Teal, L., S. Marras, M. Peck, and P. Domenici. 2018. Physiology-based modelling approaches to characterize fish habitat suitability: Their usefulness and limitations - ScienceDirect. Estunarine Coast. Shelf Sci. **201**: 56–63. doi:10.1016/j.ecss.2015.11.014

Thorson, J., N. Ianelli, S. B. Munch, K. Ono, and P. D. Spencer. 2015. Spatial delay-difference models for estimating spatiotemporal variation in juvenile production and population abundance. Can. J. Fish. Aquat. Sci. **72**: 1897–1915. doi:10.1139/cjfas-2014-0543

Thorson, J. T. 2018. Three problems with the conventional delta-model for biomass sampling data, and a computationally efficient alternative. Can. J. Fish. Aquat. Sci. **75**: 1369–1382. doi:10.1139/cjfas-2017-0266

Thorson, J. T. 2019*a*. Forecast skill for predicting distribution shifts: A retrospective experiment for marine fishes in the eastern Bering Sea. Fish Fish. **20**: 159–173. doi:10.1111/faf.12330

Thorson, J. T. 2019*b*. Guidance for decisions using the vector autoregressive spatio-temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fish. Res. **210**: 143–161. doi:10.1016/j.fishres.2018.10.013

Thorson, J. T., and L. A. K. Barnett. 2017. Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. ICES J. Mar. Sci. **74**: 1311–1321. doi:10.1093/icesjms/fsw193

Thorson, J. T., J. N. Ianelli, and S. Kotwicki. 2017. The relative influence of temperature and size-structure on fish distribution shifts: A case-study on Walleye pollock in the Bering Sea. Fish Fish. **18**: 1073–1084. doi:10.1111/faf.12225

Wood, S. N. 2006. Generalized additive models: An introduction with R, 1st ed. Chapman and Hall/CRC Press.

Zhang, J., R. Woodgate, and S. Mangiameli. 2012. Towards seasonal prediction of the distribution and extent of cold bottom waters on the Bering Sea shelf. Deep Sea Res. Part II Top. Stud. Oceanogr. **65–70**: 58–71. doi:10.1016/j.dsr2.2012.02.023

Acknowledgments

I thank M. Litzow, L. Rogers, and L. Ciannelli for discussions that prompted me to investigate spatially varying coefficient models, as well as sparked my interest in combining physical oceanography with species distribution models. I also thank C. Rooper, D. McGowan, N. Sibanda, and one anonymous review for comments on earlier drafts. Finally, I thank the many scientists who have contributed to the eastern Bering Sea bottom trawl survey analyzed here, and the developers of Template Model Builder, without which the R package *VAST* would not exist.

Conflict of Interest

None declared.

Submitted 12 March 2019
Revised 07 May 2019
Accepted 04 June 2019

Associate editor: Kelly Benoit-Bird