ICES Journal of Marine Science, 2022, **79**, 677–688
DOI: 10.1093/icesjms/fsac015
Advance access publication date: 17 February 2022
Original Article

ICES CIEM International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

Identifying species complexes based on spatial and temporal clustering from joint dynamic species distribution models

Kristen L. Omori [1,*] **and James T. Thorson** [2]

[1]Virginia Institute of Marine Science, William & Mary, 1375 Greate Rd, Gloucester Point, VA 23062, USA
[2]Habitat and Ecological Processes Research Program, Alaska Fisheries Science Center, NOAA , Seattle, WA, USA
[*] *Corresponding author: tel: 1-804-684-7885; e-mail:* komori@vims.edu

Data-limited species are often grouped into a species complex to simplify management. Commonalities between species that may indicate if species can be adequately managed as a complex include the following: shared habitat utilization (e.g., overlapping fine-scale spatial distribution), synchrony in abundance trends, consistent fishing pressure or gear susceptibility, or life history parameters resulting in similar productivity. Using non-target rockfish species in the Gulf of Alaska as a case study, we estimate spatial and temporal similarities among species to develop species complexes using the vector autoregressive spatio-temporal (VAST) model, which is a joint dynamic species distribution model. Species groupings are identified using Ward's hierarchical cluster analysis based on spatial and temporal species correlations. We then compare the spatial and temporal groupings with cluster analysis groupings that use exploitation and life history characteristics data. Based on the results, we conclude that there are some rockfish species that consistently group together, but the arrangement and number of clusters differ slightly depending on the data used. Developing species complexes for fisheries management requires a variety of analytical approaches including species distribution models and cluster analyses, and these can be applied across the full extent of available data sources.

Keywords: non-target species, rockfish, stock complex, spatio-temporal modeling, spatial overlap, VAST model.

Introduction

The management of non-target species, those being caught incidentally to the primary species, can be a challenge, particularly when the fisheries span a wide, heterogeneous geographic area. Non-target species typically have limited economic value, have low population densities, or are preserved as forage for other ecosystem components (e.g., Davies *et al.*, 2009). Despite potential limited economic importance, adequate management of non-target catch is necessary to maintain individual populations and to ensure overall ecosystem health, particularly when the accumulated biomass of non-target species exceeds the targeted species biomass or for species with high vulnerability to overfishing (e.g., Lewison *et al.*, 2004; Piet *et al.*, 2009; Cope *et al.*, 2011; Rezende *et al.*, 2019). Many non-target species have sparse life history information (e.g., age or length at maturity, maximum age), undocumented species-specific catch histories, or unknown spatial distributions. Additionally, fishery-independent surveys, if available, do not typically optimize sampling for non-target species. Yet, international fishery policies, particularly in the United States and the European Union, mandate catch limits on all exploited species, including data-limited species (e.g., Magnuson-Stevens Reauthorization Act of 2006, MSRA, 2007; Common Fisheries Policy, CFP, 2013).

One approach to comply with management requirements is to assess an assemblage of non-target species as a unit, also known as a species complex (Jiao *et al.*, 2009). Complexes are typically designated for species caught in a multispecies fishery where adequate data or technical support are often lacking to perform single-species stock assessments, or where gear interactions make single-species management difficult to implement (USOFR, 2009; Reuter *et al.*, 2010). While species complexes are characterized as a group of species having similar life history characteristics, susceptibility to the fishing gear, and geographic distributions (MSRA, 2007), information is typically missing to satisfy all requirements. For example, estimates of life history values (e.g., maximum age, age-at-maturity) can be sparse for data-limited species. Although recent efforts have predicted these for all fish species using taxonomic and life-history correlations (Thorson *et al.*, 2017; Thorson, 2020), these predictions are correlated within taxa and may not yield independent estimates for data-limited species. As a result, species are often partitioned into family or similar aggregations (DeMartini, 2019). Moreover, the fishery vulnerability and geographic overlap conditions for a complex can be difficult to address when the area of management spans a wide diversity of habitat and fishing gears. Ideally, species within a complex would demonstrate high spatial overlap and would sustain similar environmental and fishing pressures reflected by synchrony in temporal trends of abundance (Cope and Punt, 2009).

Understanding of spatio-temporal distributions of non-target species can help to better identify appropriate stock complex groupings. Overlapping fine-scale distributions for species in a complex indicates co-existence, which can occur when it is mutually beneficial for both (or all) individuals (e.g., schooling for increased predator protection; Morse, 1977; Parrish, 1991), species' fitness levels are comparable to one another (i.e., have similar abilities for reproductive success; Chesson and Kuang, 2008), resources are not limiting

Received: August 25, 2021. **Revised:** January 5, 2022. **Accepted:** January 17, 2022

(Neves *et al.*, 2018), or habitat complexity enables a diversity of species to cohabitate (Almany, 2004). However, understanding all the drivers that promote co-existence is a challenge (Neves *et al.*, 2018). Species belonging to a complex should fulfill similar niches in the ecosystem (i.e., niche overlap; Hutchinson, 1961), such that external forces would similarly influence the productivity and status of all species in the complex. Additionally, identifying commonality in spatial distributions can help determine the degree of similarity in harvest pressure across species. Species with similar spatial distribution are more likely to be harvested at similar rates (if they have similar susceptibility to the gear), thereby reducing the possibility of localized depletion if the complex is managed for sustainability (Jarillo *et al.*, 2018).

A common ecological tool for identifying spatial correlations and structure of biological populations are species distribution models (SDMs), which are becoming increasingly prevalent for fisheries applications (Planque *et al.*, 2011; Berger *et al.*, 2017). A critical advancement in SDMs, which has allowed a more thorough understanding of species distributions given limited observed data, has been the incorporation of spatial autocorrelation (Legendre, 1993; Dormann *et al.*, 2007). The assumptions of spatial autocorrelation (i.e., spatial covariation among locations) imply that neighboring locations are more similar than locations farther away, which enable an SDM to interpolate across space and estimate abundance for model grid cells that may not have data (Dormann *et al.*, 2007; O'Leary *et al.*, 2020). A wide variety of potential explanatory covariates can be incorporated into SDMs to improve performance when extrapolating density to areas where data are sparse. For instance, many SDMs incorporate abiotic factors as covariates in the model, including temperature, depth, sediment or bottom type, salinity, and spatially varying responses to regional oceanographic conditions (e.g., Nishida and Chen, 2004; Perry *et al.*, 2005; Godefroid *et al.*, 2019; Thorson, 2019). While abiotic factors can increase the predictive performance of presence or density estimates, there are many unknown factors that similarly influence a species' distribution. More recently, SDMs have been extended to incorporate random effects to account for unobserved or unexplained processes affecting the spatial distribution of species (Thorson *et al.*, 2015).

Identifying co-existence or asynchrony among species can help inform spatial distributions in the same way as abiotic factors, particularly when spatial abiotic factors are unknown. Joint dynamic SDMs allow for the simultaneous estimation of spatio-temporal densities for multiple species. Incorporating species associations can be particularly useful for data-limited species due to sparse observation data. Thus, identifying and incorporating species' relationships can improve spatio-temporal extrapolations and aid in estimating species distribution (Ovaskainen and Soininen, 2011; Thorson *et al.*, 2015; Thorson and Barnett, 2017). Modeling both spatial and temporal correlations for multiple species simultaneously can better delineate direct relationships among species' distributions and detect spatio-temporal changes in respect to biotic, abiotic, and anthropogenic factors (Godefroid *et al.*, 2019). Joint dynamic SDMs can generate community dynamic spatio-temporal trends in addition to individual abundance indices (Thorson *et al.*, 2016). Therefore, a single model to identify fine-scale spatial correlation along with similarities in temporal trends among multiple data-limited species can help validate species complexes.

The goal of this research is to explore a new application of joint dynamic SDMs as a tool for identifying species complexes for data-limited species. We demonstrate the approach through exploration of the spatial and temporal patterns of non-target rockfish belonging to two management sub-groups in the Gulf of Alaska: Other Rockfish and Demersal Shelf Rockfish complexes. The model accounts for unobserved spatial and temporal variation in a delta-model that estimates the unknown variables in both the encounter and positive catch given presence model components. The joint dynamic SDM model is applied to examine the spatial and temporal correlations along with overlap in distributions of the species in the non-target Gulf of Alaska rockfish complexes by modeling the species simultaneously. We then apply a clustering method to the results of the joint dynamic SDM to group species based on spatial and temporal synchrony in abundance and distribution. Finally, we compare resultant species complexes suggested by the joint dynamic SDM to species complex delineations based on harvest rates and on life history characteristics, along with recent assignments to complexes suggested by Omori *et al.* (2021). We show that identifying spatio-temporal correlations using joint dynamic SDMs can be a useful tool for identifying species complexes, which is a useful addition to the suite of analytical clustering approaches currently considered for identifying species complexes in data-limited situations.

Methods

Case study: Gulf of alaska non-target rockfish

In the Gulf of Alaska, non-target rockfishes (*Sebastes* spp.) are managed in two sub-groups, Other Rockfish and Demersal Shelf Rockfish complexes, depending on the management area. The Other Rockfish complex consists of over 20 non-target rockfishes within the Gulf of Alaska with more than half that are rarely caught (<1% of the total Other Rockfish catch). The Demersal Shelf Rockfish complex includes seven species that are a subset of the Other Rockfish complex but are managed separately in only one of the five management areas (southeastern most management area, 650; Supplementary Material Figure SM1). For the purposes of this document, "Other Rockfish" will refer to species within both complexes in the Gulf of Alaska. Members of the Other Rockfish complexes vary widely in their distribution and habitat preference, with many being at the northern limits of their distribution in the Gulf of Alaska, which typically span the U.S. West Coast from Southern California to Alaska (Love *et al.*, 2002). Rockfish tend to be found near the benthic substrate in a variety of habitats including high relief rocks, reefs, and boulders, to low relief rocky substrate and mudflats (Johnson *et al.*, 2003; Conrath *et al.*, 2019). The Other Rockfish species are typically found in depths ranging from 100 to 275 m but can be found at depths up to 800 m (Love *et al.*, 2002). Rockfish species generally have late-maturity, are long-lived, and bear live young (Love *et al.*, 2002; Beyer *et al.*, 2015). These low productivity characteristics tend to place rockfish into a high vulnerability category, requiring careful management (Cope *et al.*, 2011; Ormseth and Spencer, 2011).

The Other Rockfish species are incidentally caught in trawl and longline groundfish fisheries. Around 46% of these non-targeted species are discarded (Tribuzio and Echave, 2019) due to their low economic value (B. Fissel, AFSC, pers. comm.), relative smaller body size compared to other rockfish

Table 1. List of non-target rockfish species (*Sebastes* spp.) included in the multispecies VAST model, including the current complex assignment (OR= Other Rockfish, DSR= Demersal Shelf Rockfish in area 650 and is in the OR complex in all other management areas), the associated total number of trawl survey tows with a non-zero catch (i.e., encounters) and the total biomass from the bottom trawl survey catch for all years combined for each species.

Common Name	Scientific Name (*Sebastes*)	Complex Designation	Encounters	Total Biomass (kg)
canary	*S. pinniger*	DSR	72	3,399
harlequin	*S. variegatus*	OR	886	22,135
redbanded	*S. babcocki*	OR	1,256	6,270
redstripe	*S. proriger*	OR	369	19,913
sharpchin	*S. zacentrus*	OR	881	36,719
silvergray	*S. brevispinis*	OR	815	34,266
yelloweye	*S. ruberrimus*	DSR	304	2,034
yellowmouth	*S. reedi*	OR	54	1,331
yellowtail	*S. flavidus*	OR	74	1,651

species, and low catch rates. Most of the species in the Other Rockfish complex are caught in the Gulf of Alaska trawl fisheries, while a subgroup of rockfish is primarily caught by longline gear fisheries in rocky, high-relief habitats (Tribuzio and Echave, 2019).

Gulf of alaska fishery-independent trawl survey dataset

For our analyses, we use the National Marine Fisheries Service (NMFS) bottom trawl survey because this fishery-independent survey represents the most long-running and spatially extensive source of data on all Other Rockfish species. The NMFS trawl survey is also currently used as the basis of the Gulf of Alaska Other Rockfish complex stock assessment and management advice (Tribuzio and Echave, 2019). The NMFS bottom trawl survey, further referred to as the 'trawl survey', has collected species-specific data for the Other Rockfish species, including weights of each species caught per tow, since 1984; the Gulf of Alaska was surveyed on a triennial rotation from 1984 to 1996, followed by a biennial time scale from 1999 to the present (von Szalay and Raring, 2018). We include years from 1984 to 2019, which totals 16 surveyed years. The survey is conducted from May through August and follows a stratified random sampling design with, on average, 725 tows per year (Supplementary Material Figure SM1). The bottom trawl survey reaches depths up to 900 m depending on the year. Each tow covers on average 0.032 km^2 swept per tow. Species-specific data are entered as biomass caught per tow with area swept as an offset. Due to the extreme rarity of some bycatch species in the Other Rockfish complex, we demonstrate our modeling approach using a subset of nine rockfish with the requirements that each chosen species make up more than 1% of the total catch biomass in the survey over the entire time series (Table 1). These nine species have been consistently identified to species level throughout the time series.

Model structure

We implement a joint dynamic SDM that applies spatial dynamic factor analysis to identify spatial and temporal commonality among the Other Rockfish species. More specifically, we applied the Vector Autoregressive Spatio-Temporal (VAST) model version 3.4.0 (downloaded from https://github.com/James-Thorson-NOAA/VAST) developed by Thorson *et al.* (2015, 2016) to account for latent spatial and temporal variation. VAST estimates the latent (i.e., unknown) variables as "factors", and allows locations to be spatially autocorrelated and years to be independent, random, or correlated with previous time steps. This joint dynamic SDM framework allows flexibility of delta-model variants, which separates the catch into two parts: encounter probability and catch probability given presence.

We apply a Poisson-link delta-gamma model (Thorson, 2018; Thorson *et al.*, 2021) to estimate the biomass-density of each species at every location in each year. The model estimates two linear predictors for every grid cell g that are used to map densities and predict total abundance across the domain of the survey. The first linear predictor, p_1, is transformed using a complementary log-log link to predict encounter probability r_1, while the second, p_2, is transformed to positive catch rate r_2 such that expected density $\mathbb{E}(B) = a \times \exp(p_1 + p_2) = r_1 \times r_2$. These "Poisson-link delta model" transformations ensure that both linear predictors have an additive effect on log-density, and we use this property in the following sections; see Thorson (2019) and Supplementary Materials SM1 for more details. Separate linear predictors are assumed for the encounter (p_1) and positive catch (p_2) distributions, and corresponding subscripts, 1 and 2, are removed for brevity.

Specifically, the linear predictor $p(s, c, t)$ at the location s, species c, and year t is modeled as:

$$p(s,c,t) = \underbrace{\mu_\beta(c) + \sum_{f=1}^{n_\beta} L_\beta(c,f)\,\beta(t,f)}_{\textit{temporal variation}} + \underbrace{\sum_{f=1}^{n_\omega} L_\omega(c,f)\,\omega(s,f)}_{\textit{spatial variation}} \quad (1)$$

Here μ_β is an intercept parameter for each species. The temporal variation is represented by $\beta(t, f)$ for each year and factor, f (latent variable), with the associated loadings matrix denoted as $\mathbf{L}_\beta$ containing elements $L_\beta(c, f)$. The values in this loadings matrix represent the linear association between each species c and each factor f, such that the loadings matrix $\mathbf{L}_\beta$ times its transpose is equal to the covariance among species resulting from those factors. For identifiability we specify that the upper-triangle elements of each loadings matrix are fixed at 0 (Zuur *et al.*, 2003), and then subsequently rotate the matrix prior to interpretation (as discussed in detail later).

Here, the temporal variation follows a random walk:

$$\beta(t,f) = \begin{cases} \delta(t,f) & \text{if } t = t_{min} \\ \beta(t-1,f) + \delta(t,f) & \text{if } t > t_{min} \end{cases}, \quad (2)$$

where $\delta(t, f)$ follows a standard normal distribution. The $\omega(s, f)$ is the spatial variation for each location and factor,

and $L_\omega(c, f)$ components are the loadings matrices. We use a full rank model by defining the number of factors in the model as the total number of species, such that $n_\beta = n_\omega = n_c = 9$, for both the temporal and spatial variation in each linear predictor.

The spatial variation (ω) is estimated as a Gaussian Markov Random Field to account for spatial autocorrelation:

$$\omega(s, f) \sim MVN(0, \mathbf{R}) \tag{3}$$

$$\text{where } R(s_n, s_m) = \frac{1}{2^{\nu-1}\Gamma(\nu)} \times \left(\kappa \left|d(s_n, s_m)\right|\right)^{\phi} \times K_\nu\left(\kappa \left|d(s_n, s_m)\right|\right). \tag{4}$$

The spatial correlation matrix, **R**, is modeled using a Matérn correlation function, which assumes that nearby locations are more correlated and the correlation decreases by distance (Thorson, 2019). In the Matérn correlation function, ν is the smoothness parameter, where we assume $\nu = 1$, and κ represents the decorrelation rate given the distance between any two locations ($d(s_n, s_m)$).

The predicted density, $d(s, c, t)$, is estimated using a Poisson-link that assumes that areas with a higher expected encounter rate also have a higher expected biomass for each encounter. The predicted density from the Poisson-link model for each observation using the first and second linear predictors (Eq. 1) is

$$d(s, c, t) = r_1(s, c, t) \times r_2(s, c, t) \tag{5}$$

$$\text{Where } r_1(s, c, t) = 1 - \exp(-a \times \exp(p_1(s, c, t))) \tag{6}$$

$$\text{and } r_2(s, c, t) = \frac{a \times \exp(p_1)}{r_1(s, c, t)} \times \exp(p_2(s, c, t)) \tag{7}$$

The density is based on the predicted biomass (mt) per area-swept, a (km^2), where the area-swept is included as an offset in the model in Eqs. 6 and 7. The first linear predictor is transformed to r_1 using a complementary log-log link function (Eq. 6), while both linear predictors are combined to calculate r_2. The Poisson-link delta model specifies that density is calculated as the log-linked sum of the two linear predictors, such that estimates of covariance can be combined across predictors (as described in Estimated and Derived Quantities section below).

The encounter and positive catch probabilities are modeled in the delta-gamma model to obtain the probability distribution of biomass catch as

$$\Pr(b(s, c, t) = B) = \begin{cases} 1 - r_1(s, c, t) & if\ B = 0 \\ r_1(s, c, t) \times Gamma\left\{ B \,|k = \frac{1}{\sigma_c^2}, \lambda = r_2(s, c, t) * \sigma_c^2 \right\} & if\ B > 0 \end{cases}, \tag{8}$$

where the observed biomass catch data ($b(s, c, t)$) is for each location s, species c, and year t. The positive catch is modeled using a reparameterized gamma distribution for the probability density function, where the shape, k, and scale, λ, are functions of the expected catch given a presence, $r_2(s, c, t)$ (Eq. 7), and residual biomass sampling variation, σ_c^2, for species c (i.e., measurement error).

Model diagnostics that are used to determine model fit include the comparison of empirical distributions to simulated distribution using Quantile-Quantile (Q-Q) plots and density histograms (Supplementary Material Figures SM3 and SM4). The Q-Q plots are generated by calculating the quantile for each positive catch relative to the gamma distribution used to fit positive catches, and then plotting these quantiles relative to a normal distribution. The Q-Q plot indicates reasonable model fit for the positive component of the delta model if the relationship between the empirical and simulated distribution follows a linear, one-to-one line.

The flexible model structure includes fixed effects (σ_c^2, $\mu_\beta$intercepts, $L_\beta(c, f)$, $L_\omega(c, f)$) and random effects ($\beta(t, f)$ and $\omega(s, f)$). Fixed effects are estimated by maximizing the marginal likelihood while integrating across the random effects. A Laplace approximation is used to approximate the marginal likelihoods (Skaug and Fournier, 2006). To improve efficiency of estimation, VAST creates a mesh of discrete locations (i.e., knots) to represent a reduced set of locations to approximate the sampling area (Thorson *et al.*, 2015; Supplementary Material Figure SM2). The knots are determined internally in VAST by a k-means cluster algorithm, which distributes the knots based on the proportionality of sampling intensity (Shelton *et al.*, 2014; Thorson *et al.*, 2015). We use 500 knots (estimated locations, s) and the average minimum distance between knots is 20.4 km, where the geostatistical range (i.e., distance with approximately 10% correlation) for our study is estimated to be 127.1 km for the first linear predictor and 47.1 km for the second linear predictor. VAST uses a stochastic partial differential equation (SPDE) approximation for the Gaussian Markov Random Field with a Matérn correlation function for computational efficiency (Lindgren *et al.*, 2011) to estimate the spatial variation. VAST then uses bilinear interpolation to calculate the value of random fields at sampled locations, as well as at each of 23,329 extrapolation-grid cells, each representing a 2 nm by 2 nm area within the Gulf of Alaska. VAST is executed in Template Model Builder (TMB; Kristensen *et al.*, 2016) and uses R statistical program (R Core Team, 2021) as a platform to identify the maximum likelihood estimates of fixed effects. Lastly, we use Microsoft R Open 3.5.3 (https://mran.microsoft.com/) to increase computational speed via efficient and parallel computation of linear algebra. The VAST model structure is described in further detail in Thorson (2019) and the VAST code is available online (www.github.com/james-thorson-NOAA/VAST).

Estimated and derived quantities

From the estimated parameters and values from the joint dynamic SDM, we derive quantities, such as covariance matrices and indices of abundance, to examine the temporal and spatial relationships among species. The individual covariance matrices for each linear predictor are calculated using the estimated loadings matrices where the temporal loading matrices, $\hat{L}_\beta(c, f)$, and spatial loading matrices, $\hat{L}_\omega(c, f)$, are denoted as $\hat{\mathbf{L}}_1$ and $\hat{\mathbf{L}}_2$ for the first and second linear predictor. We compute a single, joint covariance matrix, $\hat{\mathbf{V}}_{\mathbf{total}}$, for each model component, temporal (β) and spatial (ω), by summing the individual covariance matrices from the first and second linear predictors as follows:

$$\hat{\mathbf{V}}_{\mathbf{total}} = \hat{\mathbf{L}}_1^{\mathrm{T}}\hat{\mathbf{L}}_1 + \hat{\mathbf{L}}_2^{\mathrm{T}}\hat{\mathbf{L}}_2. \tag{9}$$

The $\hat{\mathbf{V}}_{\mathbf{total}}$ are examined to determine temporal and spatial similarities among species by examining the positive, negative, or neutral individual variance between two species c_1 and c_2, $\hat{v}_{total}(c_1, c_2)$ in the $\hat{\mathbf{V}}_{\mathbf{total}}$. We calculate the combined loading matrices for the temporal and spatial model components

using a Cholesky decomposition on each $\hat{\mathbf{V}}_{\mathbf{total}}$ in order to examine the association of species with latent factors. To improve interpretability, we rotate each combined loadings matrix such that the first and second axes (which we subsequently plot) represent the maximum possible variance in $\hat{\mathbf{V}}_{\mathbf{total}}$. Rotations are common in factor analysis (Browne, 2001), and we specifically apply a "PCA-rotation" that is described in detail in Thorson *et al.* (2016). The proportion of variation explained by each linear predictor for spatial and temporal model components is calculated by dividing the sum of the eigenvalues from $\hat{\mathbf{L}}_1$ and $\hat{\mathbf{L}}_2$ by the sum of eigenvalues from both $\hat{\mathbf{L}}_1$ and $\hat{\mathbf{L}}_2$. The proportion of variation explained by each factor for each model component is calculated by dividing the eigenvalues associated with each factor by the sum of all eigenvalues in each model component.

We calculated the indices of abundance, $\hat{I}(c,t)$, for each species and each year, which are derived from estimates in the model for the Gulf of Alaska management area by summing all the locations, n_s, in the spatial grid:

$$\hat{I}(c,t) = \sum_{s=1}^{n_s} \left(a(s) \times \hat{d}(s,c,t) \right) \tag{10}$$

where $\hat{d}(s,c,t)$ is the predicted density in mt/km^2 at each location and is expanded by the area at each location, $a(s)$, in km^2. The derived indices of abundance are used to examine the similarities in temporal trends among species.

Cluster analyses and comparison

We apply Ward's hierarchical clustering (Ward, 1963) on the $\hat{\mathbf{V}}_{\mathbf{total}}$ to investigate species groupings based on the temporal and spatial relationships to compare with the PCA-rotation results. Then we compare the Ward's clustering results with clusters from life history information and fishery data to examine consistent groupings of species. The distances $\gamma(c_1, c_2)$ between each set of species, c_1 and c_2, used in the clustering method on the covariance matrices, $\hat{\mathbf{V}}_{\mathbf{total}}$, are calculated by

$$\hat{\gamma}(c_1,c_2) = \sqrt{\hat{v}_{total}(c_1,c_1) + \hat{v}_{total}(c_2,c_2) - 2\hat{v}_{total}(c_1,c_2)}, \tag{11}$$

using the variance of each species, $\hat{v}_{total}(c_1,c_1)$ and $\hat{v}_{total}(c_2,c_2)$, and covariance between the two species, $\hat{v}_{total}(c_1,c_2)$. A distance matrix composed of elements $\hat{\gamma}(c_1,c_2)$ is calculated for each of the temporal and spatial component using the temporal, $\hat{\mathbf{V}}_{\boldsymbol{\beta}\mathbf{total}}$, and spatial, $\hat{\mathbf{V}}_{\boldsymbol{\omega}\mathbf{total}}$, covariance matrices. We calculate the average silhouette width to determine the optimal number of clusters using R package "factoextra" (Kassambara and Mundt, 2020), where the highest value indicates the preferred number of clusters (Rousseeuw, 1987). The average silhouette width measures the similarity of objects within the same cluster compared to other clusters (i.e., examines the mean distance between all members in one cluster against the distance between members of the other clusters).

We calculate the spatial and temporal centroids for each cluster from the Ward's analysis to compare spatial and temporal trends between clusters. The individual species and location specific spatial estimates from the first, $\hat{\omega}_1(c,s)$, and second, $\hat{\omega}_2(c,s)$, linear predictors that are derived from the spatial variation component in Eq. 1 are summed to obtain a total spatial estimate for each species at each location ($\hat{\omega}_{total}(c,s) = \hat{\omega}_1(c,s) + \hat{\omega}_2(c,s)$). We average the total spatial estimates of each species belonging to the cluster to obtain the average spatial value for each cluster g, $\hat{\omega}_{total}(g,s)$. Then we map the average spatial value, $\hat{\omega}_{total}(g,s)$, for each cluster to visualize the spatial pattern associated with each cluster. The temporal centroids from the Ward's clusters are calculated differently from the spatial centroids because the temporal variation follows a random walk. First, we sum the individual temporal estimates for each species and each year from the two linear predictors ($\hat{\beta}_1(c,t)$ and $\hat{\beta}_2(c,t)$) derived from the temporal variation component in Eq. 1 to obtain the total temporal estimates, $\hat{\beta}_{total}(c,t)$. Then the difference between the total temporal estimates for each species for each year ($\Delta\hat{\beta}_{total}(c,t) = \hat{\beta}_{total}(c,t) - \hat{\beta}_{total}(c,t-1)$) for $t > t_{min}$ are calculated. Next, we average the $\Delta\hat{\beta}_{total}(c,t)$ for all the species belonging to the cluster, g, each year to obtain a $\Delta\hat{\beta}_{total}(g,t)$ as the average temporal values for each year after t_{min}. Finally, the cumulative sum of the $\Delta\hat{\beta}_{total}(g,t)$ for each cluster are calculated to transform back into the original random walk scale.

We compare the clustering results based on the spatial and temporal components of the joint dynamic SDM (i.e., VAST model) with species clusters based on life history characteristics and fisheries data (Supplementary Material Table SM2 and SM3). The comparison of life history attributes is based on values determined from a literature review of Other Rockfish species (Omori *et al.*, 2021). The values utilized in this study are as follows: age- and length-at-maturity (A_{mat} and L_{mat}, respectively), maximum age observed (as a proxy for longevity, A_{max}), and mean maximum length from the von Bertalanffy growth curve (L_∞). The resultant life history table is standardized (i.e., divided by its mean) to give equal weight to each life history characteristic before calculating the Euclidean distances (i.e., similarity among species). To compare similarities in harvest impacts, we calculate the annual harvest fraction for each species by dividing the total species-specific fisheries catch in the Gulf of Alaska by the estimated abundance (Eq. 10) for each year modeled in the joint dynamic SDM. The species-specific fisheries catch data is gathered from the NMFS Alaska Regional Office Catch Accounting System (Cahalan *et al.*, 2014) using data from 2010 to 2019, representing years when robust species-specific fisheries data have been reported. Each harvest fraction is normalized by dividing by the largest harvest fraction in the dataset, followed by calculating the Euclidean distance to determine similarity among fisheries harvest fractions across species (Supplementary Material Table SM3). Finally, we apply Ward's clustering analyses to the life history and harvest fraction distance datasets to identify species clusters, using the R package "stats" (R Core Team, 2021). Dendrograms are used to compare the relationships among rockfish included in the joint dynamic SDM from the temporal and spatial covariance matrices and the life history and harvest fraction data sources. We use the average silhouette width to determine optimal number of clusters for each data source.

Results

Spatial overlap, correlation, and clusters

Other Rockfish have varying levels of spatial overlap. Density estimates suggest that some species are found in deep waters throughout the Gulf of Alaska (i.e., harlequin and yelloweye), while others have a gulf-wide distribution, but have

densities that are more concentrated in the southeastern Gulf of Alaska (i.e., canary, yellowtail, and yellowmouth; Figure 1). The remaining species also tend to have higher densities in the southeastern Gulf of Alaska but have a more gradual decreasing density gradient moving westward (e.g., redbanded, redstripe, sharpchin, and silvergray; Figure 1).

The majority of the spatial variation is explained in the first linear predictor, encounter probability (81%). For the combined spatial component, the first three rotated factors out of nine comprise 72%, 9%, and 7% of the total spatial variation, respectively (Supplementary Material Table SM4). Factor 1 from the combined spatial component appears to be associated with both differences between the southeastern Gulf of Alaska and other areas along with distance from land. Factor 2 demonstrates a more centralized association with the southeastern Gulf of Alaska, but no distinguishable association with distance from land throughout the remaining gulf (Figure 2). Based on the PCA rotation of the spatial factors, the rockfish separate into three groups (Figure 2, see Supplementary Material Table SM5 and Figure SM5 for Factor 3 factor loadings). Two species, canary and yellowtail, which have high concentrations in the southeastern Gulf of Alaska, clustered together and separate on the Factor 2 axis. Another group consisting of harlequin, redstripe, sharpchin, and yelloweye has a small, negative Factor 2 rotated loadings and a larger Factor 1 rotated loadings (Figure 2). The four species belonging to this group tend to have higher densities throughout more areas in the gulf (e.g., higher densities between Prince William Sound and Cook Inlet; Figure 1) compared to the other two groupings. The third group with redbanded, silvergray, and yellowmouth has small, positive Factor 2 rotated loadings (Figure 2). The species in this third group have higher density concentrations in the southeastern Gulf of Alaska, but most are also found in other areas in the gulf in lower densities (Figure 1). While all rockfish demonstrate positive spatial correlation with one another, there are varying strengths of correlation (Supplementary Material Figure SM6). For example, canary and yellowtail demonstrate strong positive spatial correlation, but have weaker correlation with the other rockfish (Supplementary Material Figure SM6). Overall, the strength of the correlations among species varied but were all positive.

Ward's clustering using the spatial covariance matrix results in similar groupings as those found in the PCA rotation from the spatial component, with three suggested groupings: 1. canary and yellowtail, 2. harlequin, yelloweye, redstripe, and sharpchin, 3. redbanded, silvergray, and yellowmouth (Figure 3a). The centroid of cluster 1 (containing canary and yellowtail) has high values in southeastern Gulf of Alaska (Figure 4). The centroid of cluster 2 (harlequin, yelloweye, redstripe, and sharpchin) is elevated farther from the land, as well as waters between Prince William Sound and Cook Inlet, and Cluster 3 (redbanded, silvergray, and yellowmouth) have a higher concentration in southeastern Gulf of Alaska with more extended densities into the eastern area (Figure 4).

Temporal overlap, correlation, and clusters

The calculated indices of abundance do not appear to track one another over time (Figure 5), which is supported by the weak and wide range of positive temporal correlations among species (Supplementary Material Figure SM8). However, the abundance indices for many Other Rockfish at the end of the time series are at or above their median yearly biomass indicating a relatively stable or an increase in abundance in the more recent years (Figure 5). Redbanded, silvergray, and yelloweye abundance increases throughout the survey time series, while others show higher variability (i.e., canary and sharpchin). A few species, such as redstripe, sharpchin, and silvergray have an estimated biomass almost tenfold higher than the other species included in the model (Figure 5).

The second linear predictor explains 81% of the total temporal variation, which is the variance of a random walk process, while the first linear predictor explains only 19% of the total temporal variation. The first three rotated factor loadings out of nine of the combined temporal component account for 67%, 16%, and 12% of the total temporal variation, respectively (Supplementary Material Table SM4). Yellowtail and canary appear to separate from the other rockfish along with from one another based on Factor 1 and 2 after the PCA rotation (Figure 6; see Supplementary Material Table SM6 and Figure SM7 for Factor 3 factor loadings).

Results from Ward's clustering suggests two temporal groupings for this set of rockfish (Figure 3b). Canary and yellowtail separate into their own cluster (Cluster 1), and appear to have a decrease in temporal estimate values in the early 1990s compared to the remaining species in the other cluster (Figure 7). The average temporal values (i.e., the average yearly temporal values for the cluster on the random walk scale) of the two species cluster (Cluster 1) fluctuate more than the other cluster with seven species (Cluster 2), but both clusters appear to demonstrate a slight increase in the end of the time series where the majority of the average temporal values in the latter years are above their median value (Figure 7).

Cluster analyses comparison

The clustering on the annual harvest fractions separated the species based on levels of exploitation, with three clusters defined by high, intermediate, and low harvest fractions (Figure 3c; Supplementary Material Table SM3). Similarly, the three clusters using the life history data are divided into low, medium, and high productivity levels (Figure 3d; Supplementary Material Table SM2), where low levels of productivity correspond to large sizes at 50% maturity and maximum length, and older ages for 50% maturity and maximum age.

There are two pairs of species that are clustered together consistently for all data types (i.e., spatial overlap, temporal synchrony, harvest fractions, and life history characteristics): (i) canary and yellowtail, and (ii) silvergray and yellowmouth (Figure 3). However, these pairs of species are differentially clustered with other rockfish in two or three clusters depending on the data source.

Discussion

Our results using a set of non-target rockfish species indicate that fine-scale SDMs can be a useful tool to identify species complexes. In particular, joint dynamic SDMs (i.e., VAST, in this example) can help determine co-existence and correlation among data-limited species because shared information, particularly for infrequently caught species, improves the predictive powers of the model (e.g., Ovaskainen and Soininen, 2011; Pacifici *et al.*, 2014; Thorson *et al.*, 2015). Similarly, the application of joint dynamic SDMs can be extended to benefit cryptic species (i.e., two or more species classified as a

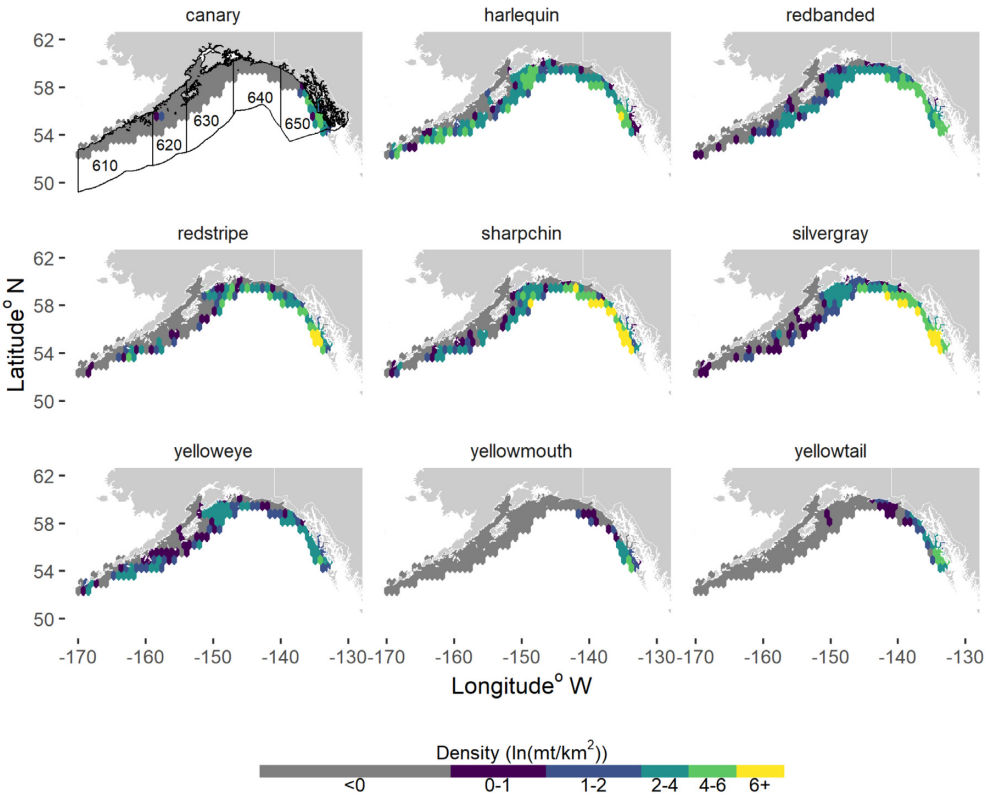

Figure 1. VAST estimated average density (ln(mt/km^2)) of each rockfish species across all Gulf of Alaska trawl survey years. The management areas are specified and denoted on the 'canary' panel.

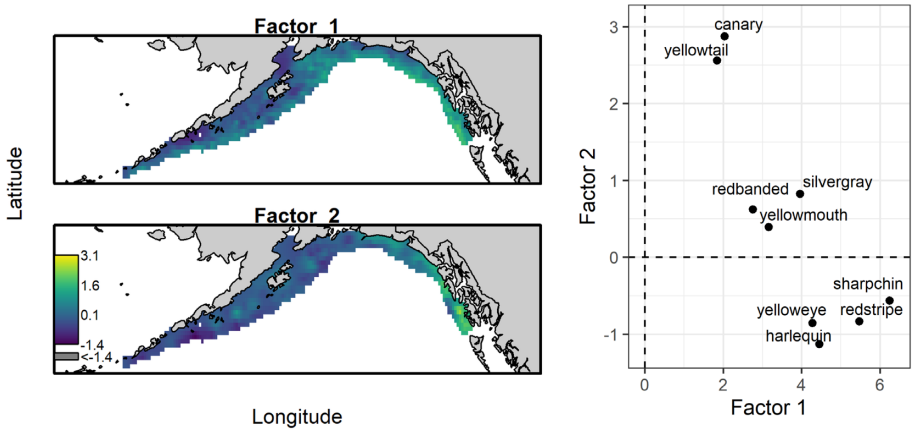

Figure 2. Combined results of the two linear predictors from the spatial component ($\omega(c,s)$) of the model. The left two panels represent maps of the first two factors after the PCA rotation. The right panel provides the first two factor loadings after a PCA rotation on the combined spatial covariance matrix for all species.

single species; Bickford *et al.*, 2007) by sharing data and assessing spatial and temporal congruence. The joint dynamic SDM applied here helps improve the understanding of fine-scale spatial distributions as well as the detection of spatial overlap across species, especially compared to using broad-scale distributions of species.

Because fine-scale joint dynamic SDMs can detect both spatial and temporal correlations among species, they can be used by managers to group species with overlapping distributions to help better understand spatial community structure. By clustering non-target species in multispecies fisheries, the application of a joint dynamic SDM can aid managers in determining incentives or appropriate regulations to decrease the fishing pressure on areas with high densities of non-target species (e.g., Dolder *et al.*, 2018; Stock *et al.*, 2020). Moreover, major changes in fishing practices or large environmental perturbations can potentially be detected across the community of a species complex or may be detected early if population trajectories of particularly susceptible species rapidly alter (Pollock *et al.*, 2014). For example, the model can help identify population shifts in the center-of-gravity (Thorson and Barnett, 2017), which will become increasingly important as more species continue to move northward (Pinsky *et al.*, 2013). However, large differences in the

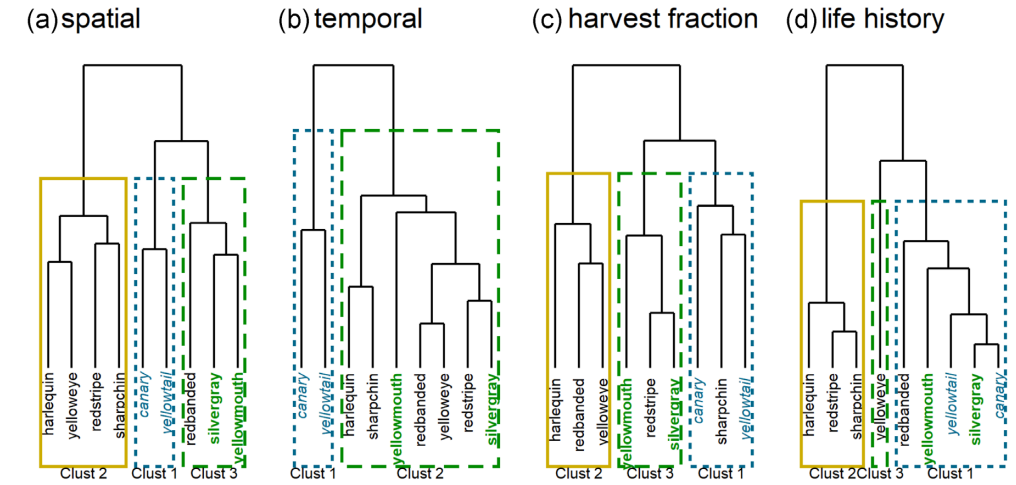

Figure 3. Ward's hierarchical clustering dendrograms from the spatial component (a), temporal component (b), harvest fraction (c), and life history (d). Box colors and line type indicate different cluster groupings with labeled cluster below, where each data type are independent from one another. Italicized species in blue and bold species in green indicate two sets of rockfish that consistently group together.

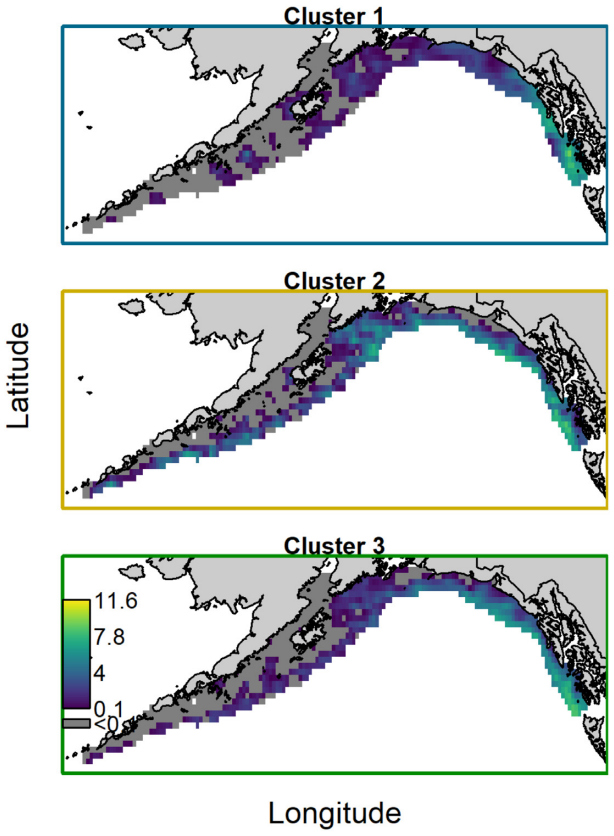

Figure 4. Average distribution of the spatial component clusters identified by the Ward's clustering results (Figure 3a, $\hat{\omega}_{total}(g, s)$). Values that are less than 1% of the maximum value are represented by dark gray and box colors correspond with spatial clusters from Figure 3a.

individual species' biomass in a complex can lead to one species component becoming overfished when large discrepancies in population sizes exist (PFMC, 2013). Careful consideration and evaluation when developing complexes is warranted, particularly for species, such as rockfish, that are vulnerable to overfishing (Cope *et al.*, 2011; Ormseth and Spencer, 2011).

The results of a joint dynamic SDM applied to fishery-independent trawl data for Gulf of Alaska rockfish species indicate that spatial factors are a key element linking these species together. The model gave evidence that the density estimates for most species is highest in the southeastern area of the Gulf of Alaska. The non-target rockfish species analyzed in this study cluster by spatial density, with three main spatial patterns: high-density concentration in only the southeastern Gulf of Alaska; high density in the eastern Gulf of Alaska but still encountered throughout the gulf; and ubiquitously distributed throughout the gulf. We did not see strong temporal correlations among species, suggesting that these rockfish demonstrate differential responses to environmental and fishing pressures. We note that no single species demonstrated a severe decrease in abundance during the time series, although harlequin decreased initially then exhibited stable biomass throughout the rest of the time series. Excluding the first few years of the survey data when there were small changes in the survey design (von Szalay and Raring, 2018) can alter the inference of the time series. If we excluded the first few years of the survey in this study, the estimated biomass for some rockfish species would be increasing, while others would appear to be stabilized. As noted, differential responses to perturbation can help stabilize a complex as a whole, which might be ideal for a group of species known for their longevity and late maturation (Love *et al.*, 2002; Ormseth and Spencer, 2011).

There are a few species that consistently clustered together based on the spatial and temporal correlation, harvest fractions, and life history characteristics: (i) canary and yellowtail; and (ii) silvergray and yellowmouth. However, all the rockfish did not group into the same distinct clusters for all data sources. In particular, most of the rockfish grouped in similar clusters except in either the harvest fractions or life history clusters. For instance, sharpchin, harlequin, and redstripe grouped together in all, but the harvest fraction clusters, whereas redbanded moved between groupings. As suggested in Cope *et al.* (2011), a hierarchical, step-wise grouping method can provide a way to assign rockfish to complexes by identifying important attributes (e.g., spatio-temporal overlap) that can be first used to separate the species. Subsequently, the groups can be sub-divided utilizing other factors (e.g.,

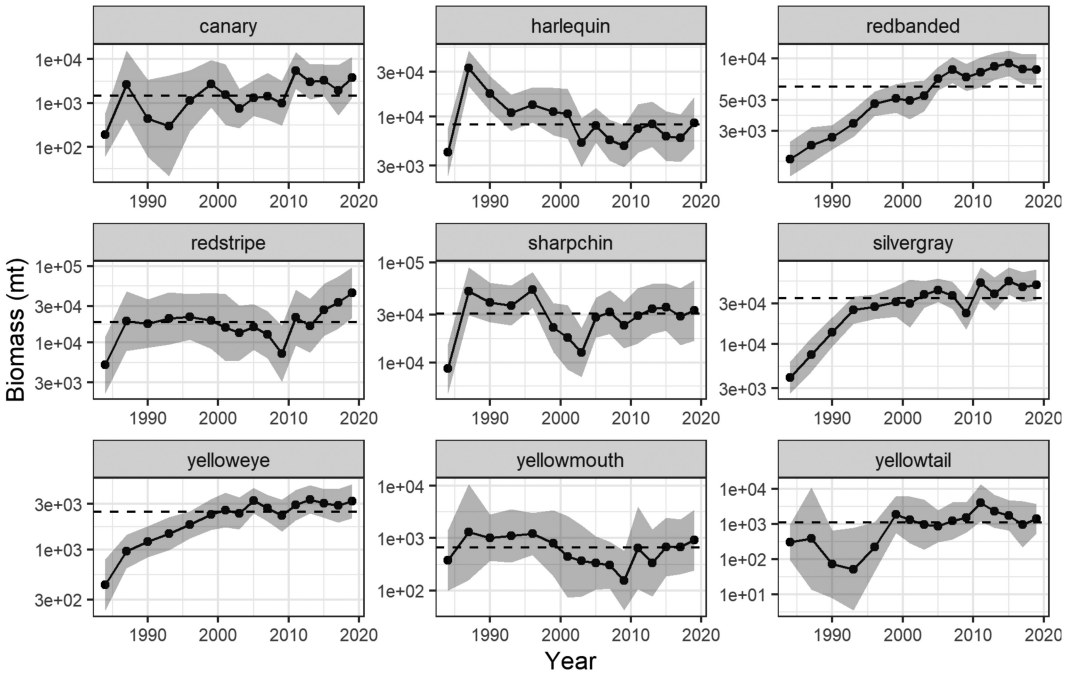

Figure 5. VAST estimated indices of abundance (black line and points) with 95% confidence interval (gray shading) and median estimate (dashed horizontal line) for non-target rockfish species.

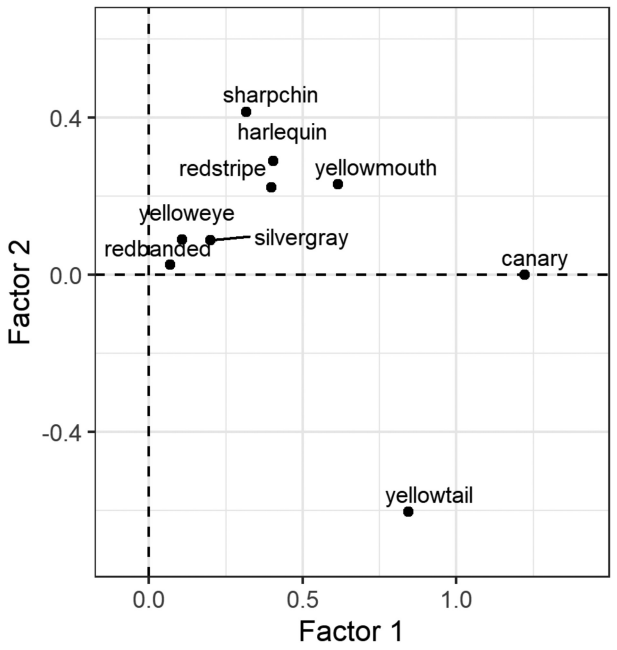

Figure 6. The first two factor loadings of the temporal variation component, $\beta(c,t)$, after a PCA rotation on the combined covariance from the first and second linear predictors, where the temporal variation follows a random walk.

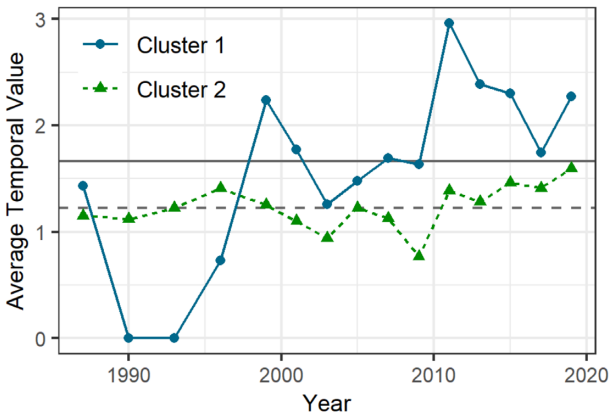

Figure 7. Average temporal values ($\hat{\beta}_{total}(g,t)$) for each cluster based on the Ward's clustering results.

fishing susceptibility or productivity) rather than comparing all variables concurrently.

Omori *et al.* (2021) performed a more broad-scale (i.e., based on management area reporting) clustering analysis to identify Gulf of Alaska rockfish groupings, which combined several surveys and fishery catch datasets, and included a different subset of rockfish due to the additional datasets. Despite differences in methodology and included species, a handful of the spatial groupings from the current fine-scale SDM are also identified in the broad-scale analysis. The most interesting results from the current study and those by Omori *et al.* (2021) regard the treatment of the Demersal Rockfish Complex managed exclusively in Gulf of Alaska management area 650. In this current study, we only included two species that belong to the Demersal Shelf Rockfish complex, canary, and yelloweye, because the remaining species in the complex are not caught in high enough numbers in the trawl survey to be modeled with a fine-scale SDM. Yet, both studies suggested that the Demersal Shelf Rockfish species should be separated from the Other Rockfish for the entirety of the Gulf of Alaska instead of only in management area 650. Yelloweye grouped with the other non-target rockfish commonly caught in the trawl survey gear. However, yelloweye is assigned with the Demersal Shelf Rockfish because it constitutes the majority of the catch for the Demersal Shelf Rockfish group, despite being caught

and distributed Gulf-wide (Tribuzio and Echave, 2019). We suggest that yelloweye continue to be managed with the Demersal Shelf Rockfish complex, but note that the other species in the complex have a smaller habitat range. The joint dynamic SDM results suggest both fine-scale spatial and temporal differences for canary and yellowtail compared to the other species in the model. Mainly, canary and yellowtail rockfish are concentrated primarily in the southeastern Gulf of Alaska with a few other patches of higher density, whereas yelloweye was spread evenly throughout the gulf. While yellowtail consistently grouped with canary in the broad-scale clustering as well as the fine-scale model, yellowtail should be carefully examined for placement into a complex because it is not considered a Demersal Shelf Rockfish.

Future spatial and temporal relationships could likely be better detected with improvements to the fine-scale SDM by incorporating other surveys that cover a wider breadth of habitat, including untrawlable areas. The trawl survey catches select species associated with the trawlable areas, which excludes habitat with high complexity. As a result, the trawl survey does not adequately sample many of the Demersal Shelf Rockfish species and other rockfish species that are associated with complex habitat (e.g., harlequin; Rooper and Martin, 2012). Further, the trawl survey depths can be restricted such that depths greater than 500 m are not surveyed in all years. Including the NMFS fishery-independent longline survey (Siwicke *et al.*, 2021) and the International Pacific Halibut Commission longline survey (Erikson and Ualesi, 2020), for instance, could extend the surveyed habitat to cover areas with increased sloping gradient and rocky habitats. The longline surveys sample a different community of rockfish species, including more Demersal Shelf Rockfish species. More surveyed habitat and different gear selectivity may help confirm the strength of the spatial correlations among non-target rockfish and increase the spatial estimation extent. Additionally, VAST has the ability to include habitat covariates in the model to help improve the density estimates as well as determine the amount of variation associated with the covariates. Rockfish are often associated with a mix of habitat types including high relief rocks, reefs, and crevices, to mudflats and vegetative areas (Johnson *et al.*, 2003; Conrath *et al.*, 2019). Adding habitat covariates, such as rocky habitat, substrate type, or depth, would help identify key attributes that influence spatial overlap of rockfish species.

Our results highlight that, when survey data are available, fine-scale SDMs can be applied to validate or construct species complexes. We demonstrate how SDMs can be used to examine both spatial and temporal similarities among species to detect fine-scale species distribution overlap and asynchronous or synchronous changes in abundance. Modeling multiple data-limited or rare species simultaneously can detect fine-scale, species-specific relationships. In comparison, multivariate approaches can utilize a wider variety of data, but typically at a broader spatio-temporal scale. Thus, multivariate approaches can provide a more general overview of potential species complexes. Joint dynamic SDMs can also help detect individual and community responses to environmental or anthropogenic perturbations, and can be used to predict how the complex may be impacted by future shifts in the ecosystem (Ovaskainen and Soininen, 2011). As species distributions continue to shift, fine-scale SDMs can aid in detecting changes in correlations among species and major shifts in their distributions. Modeling species in a complex simultaneously when species-specific data are available can help scientists provide improved management advice with limited data. In the development and management of species complexes, we advise simultaneously applying both fine-scale SDMs and broad-scale multivariate modeling techniques (e.g., Omori *et al.*, 2021), applied across the full extent of available data, to validate and/or create species complexes. Additionally, a hierarchical, step-wise structure can be used to assign species to complexes by identifying regional influential factors to separate species (Cope *et al.*, 2011). By applying the full complement of methods, including the joint dynamic SDM approach suggested here, there is greater likelihood to detect a variety of species relationships. Similarly, strong species correlations are likely to persistently appear across multiple methods and data sources, allowing the identification and validation of more robust species complexes.

Supplementary material

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Data availability statement

The data underlying this article are publicly available at https://www.fisheries.noaa.gov/alaska/commercial-fishing/alaska-groundfish-bottom-trawl-survey-data#northern-bering-sea-shelf.

Author's Contributions

K.L.O. and J.T.T. conceived the idea and designed the methodology. K.L.O. applied and adapted the model with assistance from J.T.T. K.L.O. led the writing of the manuscript with J.T.T. providing text, comments, and revisions to the manuscript.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding

KO received Federal funds under award NMFS-Sea Grant Population Dynamics Fellowship (Award NA18OAR4170322), Virginia Sea Grant College Program Project [VASG project 721691], from the National Oceanic and Atmospheric Administration's (NOAA) National Sea Grant College Program, U.S. Department of Commerce.

Acknowledgments

We express our gratitude to AFSC Bottom Trawl Survey Team for collecting the data, and the database management team for making the data assessible. We also thank Daniel Goethel, Cindy Tribuzio, Jason Cope, John Hoenig, Elizabeth Babcock, Andrew Scheld, David Johnson, and anonymous reviewers for providing helpful comments and reviews on the manuscript. This article is Contribution No. 4074 of the Virginia Institute of Marine Science (VIMS), William & Mary.

References

Almany, G. R. 2004. Does increased habitat complexity reduce predation and competition in coral reef fish assemblages? Oikos, 106: 275–284.

Berger, A. M., Goethel, D. R., Lynch, P. D., Quinn, T., Mormede, S., McKenzie, J., and Dunn, A. 2017. Space oddity: the mission for spatial integration. Canadian Journal of Fisheries and Aquatic Sciences, 74: 1698–1716.

Beyer, S. G., Sogard, S. M., Harvey, C. J., and Field, J. C. 2015. Variability in rockfish (*Sebastes* spp.) fecundity: species contrasts, maternal size effects, and spatial differences. Environmental Biology of Fishes, 98: 81–100.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng., P. K. L., Meier, R., Winker, K., Ingram, K. K., and Das, I. 2007. Cryptic species as a window on diversity and conservation. Trends in Ecology & Evolution, 22: 148–155.

Browne, M.W. 2001. An overview of analytic rotation in exploratory factor analysis. Multivariate Behavioral Research, 36: 111–150.

Cahalan, J., Gasper, J., and Mondragon, J. 2014. Catch sampling and estimation in the federal groundfish fisheries off Alaska, 2015 edition. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-AFSC-286, 46p. https://repository.library.noaa.gov/view/noaa/4833.

CFP, 2013. Regulation (EU) No 1380/2013 of the European Parliament and of the Council of 11 December 2013 on the Common Fisheries Policy. Official Journal of the European Union, 354: 22–61.

Chesson, P., and Kuang, J. J. 2008. The interaction between predation and competition. Nature, 456: 235–238.

Conrath, C. L., Rooper, C. N., Wilborn, R. E., Knoth, B. A., and Jones, D. T. 2019. Seasonal habitat use and community structure of rockfishes in the Gulf of Alaska. Fisheries Research, 219: 105331, doi: 10.1016/j.fishres.2019.105331.

Cope, J. M., DeVore, J., Dick, E. J., Ames, K., Budrick, J., Erickson, D. L., Grebel, J. *et al.* 2011. An approach to defining stock complexes for US West Coast groundfishes using vulnerabilities and ecological distributions. North American Journal of Fisheries Management, 31: 589–604.

Cope, J. M., and Punt, A. E. 2009. Drawing the lines: resolving fishery management units with simple fisheries data. Canadian Journal of Fisheries and Aquatic Sciences, 66: 1256–1273.

Davies, R. W. D., Cripps, S. J., Nickson, A., and Porter, G. 2009. Defining and estimating global marine fisheries bycatch. Marine Policy, 33: 661–672.

DeMartini, E. E. 2019. Hazards of managing disparate species as a pooled complex: A general problem illustrated by two contrasting examples from Hawaii. Fish and Fisheries, 20: 1246–1259.

Dolder, P. J., Thorson, J. T., and Minto, C. 2018. Spatial separation of catches in highly mixed fisheries. Scientific Reports, 8: 1–11.

Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G. *et al.* 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography, 30: 609–628.

Erikson, L., and Ualesi, K. 2020. IPHC Fishery-Independent Setline Survey (FISS) design and implementation in 2020. IPHC-2021-AM097-06, 12 p. International Pacific Halibut Commission, 2320 West Commodore Way, Ste 300, Seattle, WA 98199.

Godefroid, M., Boldt, J.L., Thorson, J.T., Forrest, R., Gauthier, S., Flostrand, L., Perry, R.I., Ross, A.R., and Galbraith, M. 2019. Spatiotemporal models provide new insights on the biotic and abiotic drivers shaping Pacific herring (*Clupea pallasi*) distribution. Progress in Oceanography, 178: 102198.

Hutchinson, G. E. 1961. The paradox of the plankton. The American Naturalist, 95: 137–145.

Jarillo, J., Sæther, B.-E., Enden, S., and Cao, F. J. 2018. Spatial scales of population synchrony of two competing species: effects of harvesting and strength of competition. Oikos, 127: 1459–1470.

Jiao, Y., Hayes, C., and Cortes, E., 2009. Hierarchical Bayesian approach for population dynamics modeling of fish complexes without species-specific data. ICES Journal of Marine Science, 66: 367–377.

Johnson, S. W., Murphy, M. L., and Csepp, D. J. 2003. Distribution, habitat, and behavior of rockfishes, *Sebastes* spp., in nearshore waters of southeastern Alaska: Observations from a remotely operated vehicle. Environmental Biology of Fishes, 66: 259–270.

Kassambara, A., and Mundt, F. 2020. Factorextra: Extract and visualize the results of multivariate data analyses. R package version 1.0.7. https://CRAN.R-project.org/package=factoextra.

Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. 2016. TMB: automatic differentiation and Laplace approximation. Journal of Statistical Software, 70: 1–20.

Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? Ecology, 74: 1659–1673.

Lewison, R. L., Crowder, L. B., Read, A. J., and Freeman, S. A. 2004. Understanding impacts of fisheries bycatch on marine megafauna. Trends in Ecology & Evolution, 19: 598–604.

Lindgren, F., Rue, H., and Lindstrom, J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73: 423–498.

Love, M. S., Yoklavich, M., and Thorsteinson, L. 2002. The rockfishes of the northeast Pacific. University of California Press, Berkeley, CA.

Morse, D. H. 1977. Feeding behavior and predator avoidance in heterospecific groups. Bioscience, 27: 332–339.

MSRA (Magnuson-Stevens Fishery Conservation and Management Reauthorization Act of 2006). 2007. Public Law no. 109–479, 120 Stat. 3575.

Neves, M. P., Da Silva, J. C., Baumgartner, D., Baumgartner, G., and Delariva, R. L. 2018. Is resource partitioning the key? The role of intra-interspecific variation in coexistence among five small endemic fish species (Characidae) in subtropical rivers. Journal of Fish Biology, 93: 238–249.

Nishida, T., and Chen, D. G. 2004. Incorporating spatial autocorrelation into the general linear model with an application to the yellowfin tuna (*Thunnus albacares*) longline CPUE data. Fisheries Research, 70: 265–274.

O'Leary, C. A., Thorson, J. T., Ianelli, J. N., and Kotwicki, S. 2020. Adapting to climate-driven distribution shifts using model-based indices and age composition from multiple surveys in the walleye pollock (*Gadus chalcogrammus*) stock assessment. Fisheries Oceanography, 29: 541–557.

Omori, K. L., Tribuzio, C., Babcock, E., and Hoenig, J. M. 2021. Methods for identifying species complexes using a novel suite of multivariate approaches and multiple data sources: a case study with Gulf of Alaska rockfish. Frontiers in Marine Science. 8, doi: 10.3389/fmars.2021.663375.

Ormseth, O. A., and Spencer, P. D. 2011. An assessment of vulnerability in Alaska groundfish. Fisheries Research, 112: 127–133.

Ovaskainen, O., and Soininen, J. 2011. Making more out of sparse data: hierarchical modeling of species communities. Ecology, 92: 289–295.

Pacific Fishery Management Council (PFMC). 2013. Initial Proposal (Proposed Action, Alternatives, and Considerations) for Restructuring Groundfish Stock Complexes. Agenda item D.3.a., April 2013. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101. Portland, OR, pp. 198–233. (Assessed on 25 May 2021) at: https://www.pcouncil.org/documents/2013/04/d-groundfish-management-april-2013.pdf/.

Pacifici, K., Zipkin, E. F., Collazo, J. A., Irizarry, J. I., and DeWan, A. 2014. Guidelines for a priori grouping of species in hierarchical community models. Ecology and Evolution, 4: 877–888.

Parrish, J. K. 1991. Do predators 'shape' fish schools: interactions between predators and their schooling prey. Netherlands Journal of Zoology, 42: 358–370.

Perry, A. L., Low, P. J., Ellis, J. R., and Reynolds, J. D. 2005. Climate change and distribution shifts in marine fishes. Science, 308: 1912–1915.

Piet, G. J., Van Hal, R., and Greenstreet, S. P. R. 2009. Modelling the direct impact of bottom trawling on the North Sea fish community to derive estimates of fishing mortality for non-target fish species. ICES Journal of Marine Science, 66: 1985–1998.

Pinsky, M. L., Worm, B., Fogarty, M. J., Sarmiento, J. L., and Levin, S. A., 2013. Marine taxa track local climate velocities. Science, 341: 1239–1242.

Planque, B., Loots, C., Petitgas, P., Lindstrøm, U. L. F., and Vaz, S., 2011. Understanding what controls the spatial distribution of fish populations using a multi-model approach. Fisheries Oceanography, 20: 1–17.

Pollock, L. J., Tingley, R., Morris, W. K., Golding, N., O'Hara, R. B., Parris, K. M., Vesk, P. A., and McCarthy, M. A., 2014. Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). Methods in Ecology and Evolution, 5: 397–406.

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

Reuter, R. F., Conners, M. E., Dicosimo, J., Gaichas, S., Ormseth, O., and Tenbrink, T. T.. 2010. Managing non-target, data-poor species using catch limits: lessons from the Alaskan groundfish fishery. Fisheries Management and Ecology, 17: 323–335.

Rezende, G. A., Rufener, M. C., Ortega, I., Ruas, V. M., and Dumont, L. F. C. 2019. Modelling the spatio-temporal bycatch dynamics in an estuarine small-scale shrimp trawl fishery. Fisheries Research, 219: 105336. doi: 10.1016/j.fishres.2019.105336.

Rooper, C. N., and Martin M.H. 2012. Comparison of habitat-based indices of abundance with fishery-independent biomass estimates from bottom trawl surveys. Fishery Bulletin, 110: 21–35.

Rousseeuw, P. J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20: 53–65.

Shelton, A. O., Thorson, J. T., Ward, E. J., and Feist, B. E. 2014. Spatial semiparametric models improve estimates of species abundance and distribution. Canadian Journal of Fisheries and Aquatic Sciences, 71: 1655–1666.

Siwicke, K. P., Malecha, P., and Rodgveller, C. 2021. The 2020 longline survey of the Gulf of Alaska and eastern Aleutian Islands on the FV Alaskan Leader. Cruise Report AL-20-01. AFSC Processed Rep. 2021-02, 33 p. Auke Bay Laboratories, Alaska Fisheries Science Center, NOAA, National Marine Fisheries Service, 17109 Point Lena Loop Road, Juneau, AK 99801.

Skaug, H. J., and Fournier, D. A. 2006. Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. Computational Statistics & Data Analysis, 51: 699–709.

Stock, B. C., Ward, E. J., Eguchi, T., Jannot, J. E., Thorson, J. T., Feist, B. E., and Semmens, B. X. 2020. Comparing predictions of fisheries bycatch using multiple spatiotemporal species distribution model frameworks. Canadian Journal of Fisheries and Aquatic Sciences, 77: 146–163.

Thorson, J. T., Ianelli, J. N., Larsen, E. A., Ries, L., Scheuerell, M. D., Szuwalski, C., and Zipkin, E. F. 2016. Joint dynamic species distribution models: a tool for community ordination and spatio-temporal monitoring. Global Ecology and Biogeography, 25: 1144–1158.

Thorson, J. T., Munch, S. B., Cope, J. M., and Gao, J. 2017. Predicting life history parameters for all fishes worldwide. Ecological Applications, 27: 2262–2276.

Thorson, J. T., Scheuerell, M. D., Shelton, A. O., See, K. E., Skaug, H. J., and Kristensen, K. 2015. Spatial factor analysis: a new tool for estimating joint species distributions and correlations in species range. Methods in Ecology and Evolution, 6: 627–637.

Thorson, J. T. 2018. Three problems with the conventional delta-model for biomass sampling data, and a computationally efficient alternative. Canadian Journal of Fisheries and Aquatic Sciences, 75: 1369–1382.

Thorson, J. T. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research, 210: 143–161.

Thorson, J. T. 2020. Predicting recruitment density dependence and intrinsic growth rate for all fishes worldwide using a data-integrated life-history model. Fish and Fisheries, 21: 237–251.

Thorson, J. T., and Barnett, L. A. 2017. Comparing estimates of abundance trends and distribution shifts using single-and multispecies models of fishes and biogenic habitat. ICES Journal of Marine Science, 74: 1311–1321.

Thorson, J. T., Cunningham, C. J., Jorgensen, E., Havron, A., Hulson, P. J. F., Monnahan, C. C., and von Szalay, P. 2021. The surprising sensitivity of index scale to delta-model assumptions: Recommendations for model-based index standardization. Fisheries Research, 233: 105745. doi: 10.1016/j.fishres.2020.105745.

Tribuzio, C. A., and Echave, K. B. 2019. Assessment of the Other Rockfish stock complex in the Gulf of Alaska. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska, North Pacific Fishery Management Council, 605 W 4th Ave, Suite 306. Anchorage, AK 99501. 49p.

USOFR (U.S. Office of the Federal Register). 2009. Magnuson–Stevens Act provisions; annual catch limits; national standard guidelines. Code of Federal Regulations, Title 50, Part 600. U.S. Government Printing Office, Washington, D.C.

von Szalay, P. G., and Raring, N. W. 2018. Data Report: 2017 Gulf of Alaska bottom trawl survey. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-AFSC-374, 260p.

Ward, J. H. 1963. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 58: 236–244.

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

Handling Editor: Valerio Bartolino