DR. JIE CAO (Orcid ID : 0000-0002-6250-586X)
DR. JAMES T. THORSON (Orcid ID : 0000-0001-7415-1010)
MR. CODY SZUWALSKI (Orcid ID : 0000-0002-2623-8354)

Article type : Original Article

A novel spatiotemporal stock assessment framework to better address fine-scale species distributions: development and simulation testing

Jie Cao[1,3*], James T. Thorson[2], André E. Punt[1], Cody Szuwalski[2]

1. School of Aquatic and Fishery Science, University of Washington, Seattle, WA 98195-5020, USA
2. Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, WA, United States, Seattle, WA 98112, USA

Corresponding Author Email: **jcao22@ncsu.edu**

Present address
3. Department of Applied Ecology, Center for Marine Sciences and Technology, North Carolina State University, Morehead City, NC 28557, USA

Running Head: Spatiotemporal size-structured model**ABSTRACT**

Characterizing population distribution and abundance over space and time is central to population ecology and conservation of natural populations. However, species distribution models and population dynamic models have rarely been integrated into a single modeling

framework. Consequently, fine-scale spatial heterogeneity is often ignored in resource assessments. We develop and test a novel spatiotemporal assessment framework to better address fine-scale spatial heterogeneities based on theories of fish population dynamic and spatiotemporal statistics. The spatiotemporal model links species distribution and population dynamic models within a single statistical framework that is flexible enough to permit inference for each state variable through space and time. We illustrate the model with a simulation-estimation experiment tailored to two exploited marine species: snow crab (*Chionoecetes opilio*, Oregoniidae) in the Eastern Bering Sea and northern shrimp (*Pandalus borealis*, Pandalidae) in the Gulf of Maine. These two species have different types of life history. We compare the spatiotemporal model with a spatially-aggregated model and systematically evaluate the spatiotemporal model based on simulation experiments. We show that the spatiotemporal model can recover spatial patterns in population and exploitation pressure as well as provide unbiased estimates of spatially-aggregated population quantities. The spatiotemporal model also implicitly accounts for individual movement rates, and can outperform spatially-aggregated models by accounting for time-and-size varying selectivity caused by spatial heterogeneity. We conclude that spatiotemporal modelling framework is a feasible and promising approach to address the spatial structure of natural resource populations, which is a major challenge in understanding population dynamics and conducting resource assessments and management.

Key words: fishery selectivity; Gaussian random fields; population spatial structure; spatially-explicit stock assessment model;

Table of Contents

1. INTRODUCTION

2. METHODS

2.1. Overview of methods

2.2. Model development

2.2.1. Size-structured population dynamics including spatial heterogeneity

2.2.2. Demonstration of population dynamic using two species

2.2.3. Observation models
2.2.4. Model parameters and estimation
2.3. Simulation experiments
2.3.1. Simulation overview
2.3.2. Experiment #1: Exploring movement
2.3.3. Experiment #2: Comparison of spatiotemporal and spatially-aggregated models
2.3.4. Experiment #3: Effect of sample size
3. RESULTS
3.1. Experiment #1: Exploring movement
3.2. Experiment #2: Comparing spatially-aggregated and spatiotemporal models
3.3. Experiment #3: Impact of changing sample sizes on model performance
4. DISCUSSION
ACKNOWLEDGEMENTS
DATA AVAILABILITY STATEMENT
REFERENCES
SUPPORTING INFORMATION

1. INTRODUCTION

Characterizing population distribution and abundance over space and time using mathematical and statistical models is central to population ecology and the conservation of terrestrial and aquatic organisms (Ehrlén & Morris, 2015; Krebs, 1972). These models include species distribution models (e.g., Guisan et al. 2002, Elith and Leathwick 2009) that account for abiotic and biotic covariates, and population dynamic models (Maunder & Piner, 2015) that estimate the amount of resource abundance and/or biomass (Adams, Stephenson, Dale, Ahgook, & Demma, 2008; Bieber & Ruf, 2005; Maunder & Piner, 2015). These two types of model have fundamentally different structure, so have rarely been integrated into a single modeling

framework. Consequently, natural resource management and conservation measures are often developed based on simplifying assumptions about, or implicit approximations to, population spatial structure, e.g., management of marine fisheries resources (Goethel & Berger, 2017; Goethel, Quinn, & Cadrin, 2011; Punt, Haddon, & Tuck, 2015) and terrestrial wildlife (Adams et al., 2008; Bieber & Ruf, 2005). On the other hand, studies predicting effects of environmental changes have focused primarily on species' distributions (Ehrlén & Morris, 2015). A combined approach that simultaneously estimates abundance and fine-scale spatiotemporal distribution will increase our ability to model spatially structured populations, and therefore greatly improve natural resource management and conservation.

The importance of considering population spatial structure has long been acknowledged by fisheries scientists (e.g., Beverton and Holt 1957, Berkeley et al. 2004) and terrestrial ecologists (Dunning et al., 1995; Turner et al., 1995). However, population ecology had primarily focused on developing quantitative approaches to assess resource abundance while approximating dynamics given the assumption that individuals are well mixed within the population spatial domain (i.e., spatial homogeneity), due in part to data and computational limitations. These approaches assume that population dynamics can be approximated by tracking total abundance across the entire stock, including the classical Malthusian model of exponential population growth, the Pearl-Verhulst model of logistic growth, and the Lotka-Volterra models of population interactions. Over the last two decades, investigations into population spatial structure have been at the forefront of population ecology (Ehrlén & Morris, 2015; Goethel et al., 2011; Jongejans, Skarpaas, & Shea, 2008; Punt, Haddon, Little, & Tuck, 2016; Punt et al., 2015), due in part to the lessons learned from management failures resulting from ignoring fine-scale population spatial structure (Kerr, Cadrin, & Secor, 2010). There is extensive evidence suggesting that marine and terrestrial populations are spatially patchy and locally structured (e.g., Elith and Leathwick 2009, Ehrlén and Morris 2015, Boudreau et al. 2017). In marine systems, local population processes are obscured, e.g., local depletion of weaker subpopulation or persistent high fishing pressure on local concentrations, if fine-scale population spatial structure is overlooked (Benson, Cox, & Cleary, 2015; Boudreau et al., 2017), which may lead to overexploitation of local fish populations. Locally depleted populations may not be easily replenished by recolonization (Boudreau et al., 2017; Kuo, Mandal, Yamauchi, & Hsieh, 2015). Therefore, it is critical to understand spatial population structure and address the spatial

heterogeneity in population density, productivity, and fishing pressure to prevent overfishing more vulnerable local subpopulations.

Methods have been developed to include spatial structure in assessments, either implicitly or explicitly (Punt, 2019). Models that do and do not explicitly model spatial heterogeneity are hereinafter referred to as spatially-explicit and spatially-aggregated models, respectively. However, some spatial heterogeneity can be accounted for in a spatially-aggregated model, e.g., areas-as-fleets approach, which approximates spatial heterogeneity using selectivity in a spatially-aggregated model (Berger, Jones, Zhao, & Bence, 2012; Hurtado-Ferro, Punt, & Hill, 2014). Such methods are considered as spatially-implicit. In principle, a spatially-aggregated model that matches the population biological boundaries can perform well when fishing intensity over the entire region is relatively homogeneous (Guan, Cao, Chen, & Cieri, 2013). However, this is rarely the case. Therefore, spatially-aggregated population models are likely to yield biased estimates of population quantities (Conroy, Cohen, James, Matsinos, & Maurer, 1995; Goethel, Legault, & Cadrin, 2015; Guan et al., 2013; Punt, 2019; Sampson & Scott, 2011; Turner et al., 1995), depending on the extent to which the underlying spatial structures of the population and fisheries are mis-specified (Punt et al., 2016). Spatially-explicit models have been increasingly developed to represent population spatial structure since the 1990s (Fournier, Hampton, & Sibert, 1998; Fu & Fanning, 2004; Goethel, Legault, & Cadrin, 2014; Goethel et al., 2011; Hulson, Miller, Ianelli, & Quinn, 2011; Quinn, Deriso, & Neal, 1990; Vincent, Brenden, & Bence, 2016). These models address spatial heterogeneity by dividing the region to be assessed and managed into sub-areas/subpopulations (called "spatial strata" here), within which the biological and fishery characteristics of the subpopulations are considered homogeneous, and the connectivity among strata is modeled explicitly, i.e., random (diffusive) and directed (migratory) movement of individuals among strata. Such models are hereinafter referred to as spatially-stratified models. Although these models address spatial heterogeneity to some degree, challenges and limitations remain. First, the selection of spatial strata is subject to uncertainty, depending on the understanding of the spatial population structure and data availability. Second, correlations in process errors (e.g., spatial patterns in juvenile survival) and fishing processes (e.g., spatial patterns in fishing pressure and selectivity) among spatial strata are often ignored. Third, spatially-stratified models have typically not included any spatial correlation among strata (either based on adjacency or distance). Therefore, the amount of data per stratum decreases and

147 the number of parameters increases as the number of strata increases, such that spatially-
148 stratified models have typically included a small (2-10) number of spatial strata. However, the
149 Spatial Population Model (SPM) approach of Dunn, Rasmussen, & Mormede, (2014) attempts to
150 model populations using a large number of areas.

151 Most population models that attempt to capture spatial structure in fished populations,
152 including spatially-stratified models, underuse the available spatial information because they are
153 fit to abundance index, fisheries catch, and size- or age-compositions that are aggregated
154 spatially. By doing so, population and fishery processes (e.g., density-dependence, fishing
155 pressure and selectivity) and productivity are approximated as being homogeneous across the
156 population spatial domain in spatially-aggregated models or within a stratum in spatially-
157 stratified models. Therefore, variation in survey data among sampling locations is typically
158 attributed to sampling error, while some portion of this variation actually represents predictable
159 spatial heterogeneity (e.g., Thorson and Haltuch 2018). Statistical methods and computational
160 approaches for spatiotemporal models have seen tremendous advances in recent years (Cressie,
161 Calder, Clark, Hoef, & Wikle, 2009). It is increasingly possible to fit a spatiotemporal population
162 model directly to available fishery and survey data at the scale they were collected (Boudreau et
163 al., 2017; Kristensen, Thygesen, Andersen, & Beyer, 2014; Thorson, Ianelli, Munch, Ono, &
164 Spencer, 2015). Spatiotemporal models define how population variables, e.g., density, vary
165 continuously across space (Kristensen et al., 2014), or in practice at hundreds of small-scale
166 strata, while estimating spatial variation as a random effect (Thorson et al., 2015). It would be
167 very difficult to fit spatially-stratified models with hundreds of spatial strata because each
168 individual stratum would have very little data. By contrast, the spatiotemporal approach specifies
169 that a population variable at a given location is shrunk towards estimates at a set of locations that
170 are in its neighborhood. The population variable at all locations can then be jointly estimated
171 (Kristensen et al., 2014; Thorson et al., 2015).

172 Our objective was to link species distribution and population dynamic models within a
173 single statistical framework that is flexible enough to permit inference for each state variable
174 (e.g., abundance and fishing mortality) through space and time. To do this, we build upon recent
175 research combining fish population dynamics and spatiotemporal statistics (Kristensen et al.,
176 2014; Thorson et al., 2015). In our spatiotemporal population model, we structure the population
177 by size bins because most exploited species are size-truncated, and they tend to have different

178 spatial distributions among size/age classes (Lee, Piner, Maunder, Taylor, & Methot, 2017), and
179 because these size-structured models are capable of discriminating between spatial heterogeneity
180 in recruitment and growth and survival (Thorson et al. 2015). Our model addresses the fine-scale
181 spatial structure of fish population and fisheries, which is an important challenge in
182 understanding fish population dynamics and conducting stock assessment and management.
183 We first describe the general elements of the modeling framework mathematically and
184 show how, with straightforward modifications to the population dynamic component, it can
185 accommodate a wide range of species with various types of life history. We illustrate the model
186 with a simulation-estimation experiment tailored to two exploited marine species: snow crab
187 (*Chionoecetes opilio*, Oregoniidae) in the Eastern Bering Sea and northern shrimp (*Pandalus*
188 *borealis*, Pandalidae) in the Gulf of Maine, which have different types of life history, i.e., snow
189 crab cease growth when they reach sexual maturity whereas northern shrimp do not experience
190 terminal molt. We then use theoretical and simulation examples to demonstrate model
191 performance. This includes showing that the model: 1) can recover spatial patterns in population
192 and fishing pressure and provide unbiased estimates of spatially-aggregated population quantities,
193 2) implicitly accounts for movement processes, and 3) outperforms spatially-aggregated models
194 when population density and fishing pressure are spatially heterogenous.

195 ## 2. METHODS

196 ### 2.1. Overview of methods

197 We organize our methods in two main parts, i.e., model development and simulation
198 experiments. First, we describe the estimation model (EM), including general model structure.
199 Second, we illustrate the model validation and evaluation based on simulations using the two
200 case example species.

201 ### 2.2. Model development

202 We present a size-structured spatiotemporal model, which estimates spatiotemporal
203 dynamics of size-structured populations and fisheries. To do so, we develop a model that tracks
204 variation in population density for multiple life-stages and their expected dynamics across space
205 and time. We first outline process models describing the underlying population and fishery
206 processes and illustrate the process model using the two case example species. We then specify a
207 process linking observed survey data and its associated variation to variation in population

density and fishery processes. The data model combines inference on encounter probability and abundance. We also summarize the data that are fit and the spatial scale on which they can be fit. Finally, we outline the parameters of the proposed models, and how the estimation of these parameters and other derived quantities is conducted. We represent matrices with bold uppercase notation, and vectors with bold lowercase notation. Indices used in model descriptions, data used during parameter estimation and simulation, and all parameters are listed in Table 1.

2.2.1. Size-structured population dynamics including spatial heterogeneity

We assume that population dynamics are determined by local growth and survival rates, where individuals grow from one life-stage to larger life-stages over time. Therefore, we specify a spatiotemporal size-structured population model of abundance, where $n_{s,t}(l)$ is the density (abundance, i.e. numbers per area) at location s, time t, and size-class l, and we define $\mathbf{n}_{s,t} = (n_{s,t}(1), n_{s,t}(2), \ldots, n_{s,t}(L))^T$. In general, we express the density $\mathbf{n}_{s,t+1}$ as a product of a function $g(\mathbf{n}_{s,t})$ and a process error term $e^{\boldsymbol{\varepsilon}_{s,t}}$:

$$\mathbf{n}_{s,t+1} = g(\mathbf{n}_{s,t}) \circ e^{\boldsymbol{\varepsilon}_{s,t}} \quad (1)$$

where $\mathbf{n}_{s,t}$ is a vector of densities for each of L size classes, and we use $\circ$ to indicate the elementwise product of two vectors. $g(\mathbf{n}_{s,t})$ is a potentially nonlinear function of the previous density and model parameters that describe the population dynamics. This function is general and can be chosen to match the life history of the species concerned. We demonstrate this function in detail below. $\boldsymbol{\varepsilon}_t$ is a vector of random effects that implicitly accounts for unmodeled spatial and temporal processes, e.g., movement, and spatial variation in biological parameters such as growth and natural mortality. Process errors are assumed to follow a multivariate normal distribution:

$$\mathbf{E}_t \sim \text{MVN}(0, \mathbf{R}_{spatial} \otimes \boldsymbol{\Theta}_{\boldsymbol{L}}) \quad (2)$$

where $\otimes$ denotes the Kronecker operator, $\boldsymbol{\Theta}_{\boldsymbol{L}}$ is a L by L matrix of the pairwise variance-covariance between any two size classes, and $\mathbf{R}_{spatial}$ is a Matérn correlation matrix, where the pairwise correlation between two locations s_i and $s_i + h$ is:

$$\mathbf{R}_{spatial}(s_i, s_i + h) = \frac{1}{2^{\nu-1}\Gamma(\nu)} \times (\kappa|h|)^{\nu} \times K_{\nu}(\kappa|h|) \quad (3)$$

where Γ() is the gamma function, K_ν is the modified Bessel function, κ is the parameter governing the distance h at which two locations are effectively uncorrelated, and ν is the Matérn smoothness parameter which we fix at $\nu = 1$ (Thorson & Haltuch, 2018).

2.2.2. Demonstration of population dynamic using two species

We consider two population dynamics models: one for invertebrates that exhibit an effective cessation of growth at some stage in the life history, and a general model that can be used for fish and invertebrate which do not exhibit cessation of growth. Eastern Bering sea snow crab and Gulf of Marine northern shrimp are the examples of the two population dynamic types, respectively.

Example 1: Eastern Bering sea snow crab

As a first example, we model abundance of snow crab in the Eastern Bering Sea. Snow crab are distributed on the continental shelf of the Bering Sea and are common at depths less than 200 m. The U.S pot fishery began in the 1970s after the Japanese started harvesting snow crab in the 1960s (but were subsequently excluded from the fishery in the early 1980s). The fishery peaked in the 1990s and crashed in the 2000s. This species is one of the most important crab species in terms of volume landed and value (Abbott, Garber-Yonts, & Wilen, 2010). Research has shown that spatial dynamics of the snow crab is likely affected by water temperature, sea ice extent, and other environmental factors (Mueter & Litzow, 2008; Parada, Armstrong, Ernst, Hinckley, & Orensanz, 2010).

We model the population dynamics by sex and maturity state because males and females experience different fishing mortality rates (f) (only males are retained by the fishery) and they cease growth when they reach sexual maturity. The size-specific abundance density over time is controlled by recruitment ($\mathbf{r}_{s,t}$, a vector of length *L* representing the number of juveniles per area recruiting into each size class of the modelled population), growth ($\mathbf{G}$, a sex-specific matrix describing the proportion of individuals staying in the same size class or growing into other size classes), natural mortality ($\mathbf{m}$), and fishing mortality (f). We express $g(\mathbf{n}_{s,t})$ by sex indicated by superscripts male and female as:

$$g\left(\mathbf{n}_{s,t}^{\text{male}}\right) = \begin{cases} \mathbf{r}_{s,t}p^{\text{male}} + \mathbf{G}^{\text{male}}\left(\mathbf{n}_{s,t-1}^{\text{male}} \circ \exp\left(-\mathbf{m}_{s,t-1} - \mathbf{v}f_{s,t-1}^{\text{male}}\right)\right) \circ \left(1 - \mathbf{w}^{\text{male}}\right), & \mathbf{n} = \mathbf{n}^{\lambda} \\ \mathbf{G}^{\text{male}}\left(\mathbf{n}_{s,t-1}^{\text{male}} \circ \exp\left(-\mathbf{m}_{s,t-1} - \mathbf{v}f_{s,t-1}^{\text{male}}\right)\right) \circ \mathbf{w}^{\text{male}} + \mathbf{n}_{s,t-1}^{\text{male}} \circ \exp\left(-\mathbf{m}_{s,t-1} - \mathbf{v}f_{s,t-1}^{\text{male}}\right), & \mathbf{n} = \mathbf{n}^{\omega} \end{cases} \quad (4)$$

$$g\left(\mathbf{n}_{s,t}^{\text{female}}\right) = \begin{cases} \mathbf{r}_{s,t}\left(1 - p^{\text{male}}\right) + \mathbf{G}^{\text{female}}\left(\mathbf{n}_{s,t-1}^{\text{female}} \circ \exp\left(-\mathbf{m}_{s,t-1}\right)\right) \circ \left(1 - \mathbf{w}^{\text{female}}\right), & \mathbf{n} = \mathbf{n}^{\lambda} \\ \mathbf{G}^{\text{female}}\left(\mathbf{n}_{s,t-1}^{\text{female}} \circ \exp\left(-\mathbf{m}_{s,t-1}\right)\right) \circ \mathbf{w}^{\text{female}} + \mathbf{n}_{s,t-1}^{\text{female}} \circ \exp\left(-\mathbf{m}_{s,t-1}\right), & \mathbf{n} = \mathbf{n}^{\omega} \end{cases} \quad (5)$$

where superscripts λ and ω indicate immaturity and maturity, respectively, $\mathbf{w}$ is a vector representing the proportion w_l of immature individuals at length l that mature, p^{male} is the sex ratio of the recruits, $\mathbf{m}$ is a vector of natural mortality, f is fully selected fishing mortality, and $\mathbf{v}$ is a vector of selectivity coefficients.

We complete this dynamical model by specifying how male and female abundance is initialized in the first modeled year:

$$g\left(\mathbf{n}_{s,1}^{\text{male}}\right) = \mathbf{r}_{s,1}p^{\text{male}} \circ \exp\left(\boldsymbol{\varphi}_{\text{male}}\right) \quad (6)$$

$$g\left(\mathbf{n}_{s,1}^{\text{female}}\right) = \mathbf{r}_{s,1}\left(1 - p^{\text{male}}\right) \circ \exp\left(\boldsymbol{\varphi}_{\text{female}}\right) \quad (7)$$

where $\boldsymbol{\varphi}_{\text{male}}$ and $\boldsymbol{\varphi}_{\text{female}}$ are vectors representing abundance per area at size of the first modeled year for males and females, respectively. These densities at size are assumed constant across the study area. This model assumes that females are not fished (i.e., $f_{s,t-1}^{\text{female}} = 0$) given that discard mortality of females is very low (Szuwalski & Punt, 2015).

The predicted harvest per area removed by the fishery, $\mathbf{c}_{s,t}$, for snow crab is calculated as:

$$\mathbf{c}_{s,t} = \left(1 - \exp\left(-\mathbf{v}f_{s,t}^{\text{male}}\right)\right) \circ \mathbf{n}_{s,t}^{\text{male}} \circ \exp\left(-0.5\mathbf{m}_{s,t}\right) \quad (8)$$

This catch equation implicitly assumes that fishing during year t takes place rapidly at the middle of the year because the fishery for male snow crab is considered to take place as a pulse.

Example 2: Gulf of Maine northern shrimp

As a second example, we model abundance of northern shrimp in the Gulf of Maine. The Gulf of Maine marks the southern-most extent of the species' range. Therefore, this population is sensitive to ocean temperature changes. The population is estimated at the lowest level ever and has experienced failed recruitment for the past several years (Cao, Chen, & Richards, 2017b). Consequently, the fishery has been closed since the 2014 fishing season.

Here sex is not distinguished because northern shrimp are hermaphroditic, so we specify $g(\boldsymbol{n}_{s,t})$ for both males and females as:

$$g(\mathbf{n}_{s,t}) = \mathbf{G}(\mathbf{n}_{s,t-1} \circ \exp(-\mathbf{m}_{s,t-1} - \mathbf{v}f_{s,t-1})) + \mathbf{r}_{s,t} \quad (9)$$

and initialize density using

$$g(\mathbf{n}_{s,1}) = \mathbf{r}_{s,t} \circ \exp(\boldsymbol{\varphi}) \quad (10)$$

These densities at size are assumed constant across the study area.

The predicted catch per area, $\mathbf{c}_{s,t}$, for northern shrimp is calculated using the Baranov catch equation assuming fishing mortality takes place continuously over a modeled time unit, which matches the characteristic of the fishery:

$$\mathbf{c}_{s,t} = \frac{\mathbf{v}f_{s,t}}{\mathbf{v}f_{s,t} + \mathbf{m}_{s,t}} \circ (1 - \exp(-\mathbf{m}_{s,t} - \mathbf{v}f_{s,t})) \circ \mathbf{n}_{s,t} \quad (11)$$

2.2.3. Observation models

Understanding species population dynamics usually requires ecologists to collect data using biological surveys, where sampling occurs at pre-defined sites, the amount of each species (counts or biomass) is recorded, and the sampled animals are subsampled to collect more biological information (e.g., maturity, sex, age and size). We let $x(s_i,t_i)$ represent *i*th observed count or biomass of a species sampled at a site indexed by the spatial location s_i within a study area S, during time $t_i \in \{1, ..., T\}$, and seek to specify a model relating $x(s_i,t_i)$ to predicted population density at that location and time (i.e., $\boldsymbol{n}_{s,t}$ in Eq. 1). The sampling locations are sometimes outside species' occupied habitat, so we account for encounter probability in the observation model. When fitting to samples of biomass, we represent the relationships among observed biomass at location s_i at time t_i, $x(s_i,t_i)$, predicted local density of individuals, $n(s_i,t_i)$, and encounter probability, $p(s_i,t_i)$, using a "Poisson-link" delta-model (Thorson, 2017). The encounter probability is a function of local density:

$$p_i = 1 - \exp(-a_i n(s_i,t_i)) \quad (12)$$

where a_i is the area swept for sample *i*, which is taken at location s_i and time t_i , where the number of observed individuals follows a Poisson process with expectation $n(s_i,t_i)$. Predicted positive biomass at location s_i at time t_i, $b(s_i,t_i)$, is then calculated as:

$$b(s_i,t_i) = \frac{a_i n(s_i,t_i) \times \delta(s_i,t_i)}{p_i} \quad (13)$$

where $\delta(s_i,t_i)$ is the predicted average weight of individuals.

The probability density function for biomass sample $x(s_i,t_i)$ is defined as:

$$\Pr(B = x(s_i,t_i)) = \begin{cases} 1 - p_i & \text{if } x(s_i,t_i) = 0 \\ p_i \times g(B; b(s_i,t_i), \sigma_D^2) & \text{if } x(s_i,t_i) > 0 \end{cases} \quad (14)$$

where $b(s_i,t_i)$ and σ_D^2 are the mean and variance of B and the probability density function $g(B)$ can be lognormal or gamma. This Poisson-link model ensures that biomass-sampling data can be explained via a log-linked model for population abundance $n(s,t)$ for multiple locations and times.

The observation model assumed for discrete count data is the overdispersed lognormal Poisson process:

$$\Pr(N = x_{s_i,t_i}) = \text{Poisson}(N; n(s_i,t_i) \times \exp(\eta_i)) \quad (15)$$

where x_{s_i,t_i} in this case is an observed count, and η_i is an observation-level random effect, which follows a normal distribution representing lognormal overdispersion.

Spatially-referenced fisheries-dependent data, i.e., total amount of fish by size class removed from the ocean, which are collected directly from the commercial and recreational fisheries are used to estimate the spatial pattern of fisheries. The total catch by area/size is assumed to be lognormally distributed.

2.2.4. Model parameters and estimation

The spatial variation of recruitment density, $\mathbf{r}_{s,t}$, is confounded with spatial process error (i.e., spatial variation in the density of each of the L size classes) if modeled separately (e.g., $r_{l,t}$ ~MVN(r_μ,$\mathbf{R}_{spatial}$) where r_μ is the average recruitment density at each location), because the size classes used to define recruitment are included in L. We therefore allow spatial process error to account for spatial variation in recruitment and estimate the annual average recruitment (i.e., r_t). In this case, $\mathbf{r}_{s,t}$ is equal to r_t for all locations s.

For fishery processes, log-fishing mortality at each location in each year $\log(f_{s,t})$ is modeled as a random walk process given fishing mortality the previous year:

$$\log(f_{s,t}) | \log(f_{s,t-1}) \sim \text{N}(\log(f_{s,t-1}), \sigma_f^2) \quad (16)$$

Size-specific selectivity is modeled using a logistic function of individual size, allowing the probability of capture to vary with fish size:

$$v_l = \frac{1}{1 + e^{-\theta(d_l - l_{50})}} \quad (17)$$

where θ and l_{50} are the parameters of the logistic function and d_l is the midpoint of size-class l.

The parameters that are treated as fixed effects include process error covariance ($\boldsymbol{\Theta_L}$), the parameter controlling the distance at which spatial correlations reach 10% (κ), average recruitment density for each time (r_t), initial density at size ($\boldsymbol{\varphi}$), the selectivity parameters (θ and l_{50}), and the standard deviations of fishing mortality and observations (σ_f and σ_D). We assume the growth transition matrices $\mathbf{G}$, the proportion male at recruitment p_{male}, the proportion achieving maturity at each size $\mathbf{w}$, and natural mortality rate for each size class $\mathbf{m}$ are specified based on external information (values used in the two case studies are listed in Table A1 and A2). Future work could explore estimating these parameters using additional data or meta-analytic information.

We treat the fully-selected fishing mortality at each location over time ($\mathbf{f}_s$), and the density for each size class and time ($\mathbf{N}$) as random effects. We treat density as a random effect, rather than process error ($\mathbf{E}_t$), because this state-space parameterization leads to faster parameter estimation in a similarly structured model (Thorson, Munch, & Swain, 2017). To estimate the fixed effects, we maximize the marginal likelihood function after integrating across the random effects. We use Template Model Builder, TMB (Kristensen, Nielsen, Berg, Skaug, & Bell, 2015) called from within the R statistical environment (R Core Team, 2019) to do so. The detailed procedure of estimating parameters and uncertainty using TMB is described in (Thorson, Jannot, & Somers, 2017).

Computational issues arise when modeling spatiotemporal population dynamics as Gaussian Fields (GFs). Consequently, we use a stochastic partial differential equation approximation to the GF (i.e., $\boldsymbol{\Sigma}_t$) based on a Gaussian Markov random field (GMRF) (Kristensen et al., 2015). This approach is based on a triangulation of the spatial domain, where a mesh is created based on a predefined number of nodes ('knots'). The number of knots determines the spatial resolution of the model, and is chosen as a trade-off between the accuracy of the GMRF representation and computational cost. The scale at which survey data were collected is not likely to coincide with the spatial scale of the model, i.e., the knots. Therefore, we model the abundance densities and use a "predictive process" formulation wherein we assume these function-valued variables are piecewise constant in the neighborhood of the knots.

When calculating total abundance for year t, n_t, the densities at the modeled spatial locations are scaled up by the total area associated with the knots:

$$n_t = \sum_{j=1}^{J} A_j \sum_{l=1}^{L} n_{l,j,t} \quad (18)$$

where $\mathbf{n}_{j,t}$ is the local density at knot j for each size class, A_j is the total area associated with knot j, and J is the number of knots specified. Fishery catch for each size class is aggregated at knot-level as input data. Therefore, total catch of year t, c_t, is calculated as:

$$c_t = \sum_{j=1}^{J} \sum_{l=1}^{L} c_{l,j,t} \quad (19)$$

In summary, the input data for the model are fishery-independent survey data, i.e., number of individuals or biomass by size bin, on the spatial scale which they were collected, and fishery catch for each size class aggregated at knot-level.

2.3. Simulation experiments

2.3.1. Simulation overview

We illustrate the model validation and evaluation using simulations tailored to snow crab and northern shrimp. Specifically, we conducted three simulation experiments to (1) explore how the spatiotemporal model performs when individual movement processes are modeled explicitly in the operating model (OM), (2) compare estimation performance for spatially-aggregated and spatiotemporal models, and (3) evaluate the impact of changing sample size. To do so, we developed two OMs to simulate snow crab and northern shrimp populations, respectively. The OMs have a fine spatial scale, i.e., 2 km by 2 km grid. The snow crab OM has 36,140 grid cells that represent the area surveyed in the eastern Bering Sea bottom trawl survey, and the northern shrimp OM has 4,997 grid cells over the sampling domain of the shrimp summer survey in the western Gulf of Maine. We use snow crab example to explore movement, northern shrimp example to compare models, and both species to quantify the effect of sample size.

The process model in the OM is the same as the estimation model (i.e., Eq. 1) except movement is explicitly accounted for:

$$\mathbf{N}_{t+1} = g(\mathbf{M}\mathbf{N}_t) \circ e^{\boldsymbol{\Sigma}_t} \quad (20)$$

where $\mathbf{M}$ is a movement matrix (number of grid cells by number of grid cells) representing proportions of individuals that stay at their current location or move to other locations during a given time unit t, $\mathbf{N}_t$ is an OM abundance matrix (number of grid cells by number of size classes).

Movement is assumed to occur at the start of each model time step. We derive annual **M** from instantaneous movement rates (Appendix 1). There is no movement when **M** is an identity matrix.

In general, simulating the population and fishery dynamics involves the following steps: (1) specifying the information used in the OMs (summarized in Tables A1 and A2 for snow crab and northern shrimp, respectively); (2) simulating recruitment and fishing mortality on the grids over time, $\mathbf{r}_{s,t}$ and $f_{s,t}$ (see Appendix 2 for detailed description); (3) calculating the abundance of each size class $\mathbf{n}_{s,t}$ (using Eq. 20). Thus, within one model time step, the OMs track both dynamics occurring within a single grid cell (i.e., survival, size transition and reproduction) and dynamics occurring among grid cells (i.e., movement).

Sampling processes were simulated on the grid spatial scale. For each size class, the observed catch of a randomly selected survey location and fishery removal for each grid were simulated from lognormal distribution with standard deviation of 0.3 for both species (see Appendix 3 for detailed description).

2.3.2. Experiment #1: Exploring movement

This simulation experiment uses snow crab as the example species and is designed to examine how model performance is affected by movement processes that are modeled explicitly in the OM. Additionally, we evaluated whether the spatiotemporal population model (1) captures the spatial structure and pattern in population abundance and fishing mortality of each size class over time, and (2) provides unbiased estimates of spatially-aggregated abundance and model parameters, such as fishery selectivity. We examined three scenarios: (1) the data are generated without measurement error and there is no movement in the OM; (2) same as scenario 1, except there is movement; and (3) the data have measurement error and there is movement in the OM. We only consider diffusive movement here. The instantaneous movement rate applied in OM is $0.4yr^{-1}$, i.e. about 35% of crabs within a given grid in the OM move out every year. Scenarios without measurement error are designed to check whether model is unbiased when correctly specified.

More details of the snow crab OM could be found in Appendix 1. To generate the fishery-independent survey data from the OM, 200 sites (grid cells) in the OM were randomly sampled each year. For each site, total abundance by size class and the total area of the sampled

site were recorded. Fishery catch-at-size was calculated at each grid (using Eq. 8) and then aggregated to the knot level as data for the EM. Therefore, the input data of fishery-independent and -dependent data have different spatial resolutions, i.e., grid cell and knot, respectively. For the scenarios with measurement error, we generated 100 replicated data sets with sampling errors, i.e., grid-based survey abundance and fishery catch data were assumed to be lognormally distributed. We specified 100 knots in the EM (each one representing local densities simulated within 361 grid cells on average in the OM) to approximate the fine-scale spatial processes simulated in the OM. Real-world applications would likely explore sensitivity to the number of knots, although we do not do so here.

2.3.3. Experiment #2: Comparison of spatiotemporal and spatially-aggregated models

This experiment is based on northern shrimp in the Gulf of Maine. The intent of this experiment is to demonstrate the importance of accounting for spatial processes when modeling the population dynamics of marine species. Therefore, we compare the performance of our spatiotemporal model with a spatially-aggregated model designed for northern shrimp (Cao, Chen, & Richards, 2017a). The spatially-aggregated model uses the same equations to describe the size-structured population dynamics as the spatiotemporal model. However, it can only track the spatially-aggregated size-structured population over time. Thus, it ignores all the spatial heterogeneities (i.e., all model variables are assumed to be constant across space). The technical details of the spatially-aggregated model can be found in (Cao et al., 2017a).

We simulate the size-structured population dynamics over 20 years. The simulation procedure is similar to the snow crab example, except $\mathbf{n}_{s,t}$ is calculated for both sexes simultaneously and movement is not considered. The shrimp population and fishery are simulated to mimic the real world where the stock size experienced a boom-and-bust circle (Table A2; Cao et al., 2017b). The fisheries mainly occur in inshore waters. Therefore, we simulate the fishing mortality with the spatial structure so that inshore areas have consistently higher fishing mortality over years (see Appendix 2).

Fifty simulated data sets are analyzed using the spatiotemporal and spatially-aggregated models, where spatially explicit sampling is aggregated prior to fitting in the spatially-aggregated model (following the process in Appendix 3). Each simulated data set includes survey catch rates by size, with an intensity of 1,200 sampling tows (60 sampling locations per year) and catch-at-

size in all the grid cells for 20 years. To make a fair comparison, we ensure that, for each model iteration, the data used in both estimation models are the same at the grid spatial scale (the input data for spatially-aggregated model are aggregated across grids, see Appendix 3 for details), and the values of the pre-specified life history parameters (natural mortality and growth) are the same for the spatiotemporal and spatially-aggregated estimation models. We use 50 knots for the spatiotemporal model. To assess model performance in this simulation experiment, we keep the replicates where both estimation models converged (i.e., final gradient of the likelihood <0.001 and the Hessian of fixed effects was positive definite). For converged replicates, we record all parameter estimates as well as model predictions. We also record the number of non-converged runs.

The spatiotemporal model estimates population density for each of 50 knots, while the spatially-aggregated model estimates population abundance for the entire area. Therefore, to compare results we convert results from the spatiotemporal model to a metric that is directly comparable with the spatially-aggregated model. To do so, we compare the estimates of abundance-at-size, fishing mortality at size and spawning stock biomass aggregated over the spatial domain from both models. The population-level fishing mortality f_t and aggregate selectivity-at-length $s_{l,t}$ for each size class is determined by solving:

$$c_{l,t} = (1 - \exp(-s_{l,t} f_t)) n_{l,t} \exp(-m_t) \quad (21)$$

for a given value of catch $c_{l,t}$, abundance $n_{l,t}$, and natural mortality rate m_t, where $c_{l,t} = \sum_{j=1}^{J} c_{l,j,t}$ and $n_{l,t} = \sum_{j=1}^{J} n_{l,j,t}$ are aggregated total catch and abundance for size class l and year t, respectively. After solving for $s_{l,t} f_t$ for each year t and size l, we separately identify fishing mortality f_t and selectivity-at-length $s_{l,t}$ by defining $argmax_l(s_{l,t}) = 1$ for each year t: this definition of selectivity is common in fisheries stock assessment modelling (Sampson, Scott, & Quinn, 2011). Aggregate selectivity-at-length $s_{l,t}$ in the spatial operating or estimation model is not constrained to follow any parametric shape, even though local selectivity at length v_l follows a logistic function. In particular, aggregate selectivity-at-length $s_{l,t}$ will differ from local selectivity v_l whenever fishing mortality varies strongly among spatial locations (Sampson & Scott, 2012; Sampson et al., 2011).

We finally compare the estimates with the true values and calculate the relative error in percentage for each year:

$$e_{i,k} = \left(\frac{est_{i,k} - true_{i,k}}{true_{i,k}}\right) \times 100\% \quad (22)$$

where $est_{i,k}$ and $true_{i,k}$ are estimated and true values of *i*th model and *k*th replicate. We also calculate root-mean-square error (RMSE) and relative bias (RB) in percentage of estimated abundance at size across years:

$$\mathrm{RMSE}_l = \sqrt{\frac{\Sigma_t \left(n_{l,t}^{est} - n_{l,t}^{true} / n_{l,t}^{true}\right)^2}{\tau}} \times 100\% \quad (23)$$

$$\mathrm{RB}_l = \frac{\Sigma_t \left(n_{l,t}^{est} - n_{l,t}^{true} / n_{l,t}^{true}\right)}{\tau} \times 100\% \quad (24)$$

where $n_{L,t}^{est}$ and $n_{L,t}^{true}$ are the estimated and OM abundance of size class l in year t, and τ is the number of years.

2.3.4. Experiment #3: Effect of sample size

We finally conduct a systematic simulation experiment to evaluate model performance given different sample sizes and provide some insight on the data requirements for the spatiotemporal model to have acceptable performance. We use both snow crab and northern shrimp as example species and examine three levels of sampling intensity: 50, 100, and 200 locations per year, representing data poor, moderate level, and data rich scenarios. For each scenario, we repeated the sampling process 50 times, so each iteration has different sampling locations. We examine the model convergence rate (final gradient of the likelihood <0.001) and root-mean-square error (RMSE) and relative bias (RB) in percentage of estimated abundance at size across years for each scenario.

3. RESULTS

3.1. Experiment #1: Exploring movement

The spatiotemporal model can generate unbiased and precise estimates of abundance and fishing mortality spatially when data are not subject to measurement error and no movement occurs. This confirms that the model is unbiased when correctly specified and when data are highly informative as expected from maximum likelihood theory. A comparison between simulated and estimated abundance illustrates how the spatiotemporal population model is able to reconstruct spatial variation in abundance over time (Fig. 1). The general spatial pattern of

each size class is recovered by the EM (Fig. 1). However, the estimated spatial distribution misses some fine-scale patterns, e.g., size class 3 in year 1 (Fig. 1). This is because the EM has a much coarser spatial resolution than the OM (i.e., 100 knots to approximate 36,140 grid cells) and integrates over fine-scale variation. As the number of knots increases, the estimated distribution would become smoother. The estimated spatial distribution has lower total variation (reflected by the contrast in color) than the true distribution. This is due to shrinkage, i.e., the estimate of abundance at a given location is shrunk towards the average of its neighboring locations. The spatiotemporal model is able to perfectly recover the spatial variation when the spatial scale at which data were collected and the EM operates matches (i.e., fishery catch data are aggregated to a knot-level, which matches the model spatial scale). As a result, fully-selected fishing mortality at the knot level is unbiased through space and time (Fig. A2), and selectivity is also accurately estimated (i.e., estimated θ = 0.05; l_{50} = 75.15mm; simulated θ = 0.05; l_{50} = 75mm). Accurate and precise estimates of abundance and catch are also obtained when model outputs are aggregated spatially (Fig. 2).

The model accounts for movement implicitly via its estimates of process error when the spatiotemporal model fits to data without measurement error but generated given unmodeled (in the EM) individual movement. This unmodeled spatial process did not lead to poorer model performance. The model recovers the spatial variation in abundance and fishing mortality over time (Figs. A3 and A4), and the total abundance is estimated accurately (Fig. A5). Finally, the spatiotemporal model converges (maximum gradient of the likelihood <0.001) for all 100 simulation replicates when fitted to data given process error and individual movement. The model is able to recover the spatial variation and accurately estimate the spatially-aggregated abundance and catch with lower precision (Fig. 3), when the sampling errors are present.

3.2. Experiment #2: Comparing spatially-aggregated and spatiotemporal models

The relative errors of estimated abundance-at-size suggest that the spatiotemporal model has lower bias and therefore lower errors than the spatially-aggregated model (Fig. 4). The relative errors of the spatiotemporal model fluctuate around zero across years, where some of the years have relative errors centered around zero and others have the median relative error below and above zero (Fig. 4; the median RMSE across years are about 4% for all five size classes, and the median RB are negative for the first three size classes (-1.99%, -0.99%, -0.19%) and positive

for the other two, i.e., 1.77%, 0.26%). The spatiotemporal model always underestimates the abundance of all size-classes for years 9 and 18. However, the spatially-aggregated model produced biased estimates of abundance for all size classes as none of the relative errors are centered around zero (the median RB are -5.93%, -5.65%, -5.35%, -5.98%, and -11.27% for the five size classes, respectively). Almost all of the relative errors are below zero, suggesting that the spatially-aggregated model consistently underestimates abundance-at-size. Consequently, the spatially-aggregated model underestimates the total abundance and spawning stock biomass for all years by about 10 to 20% (Fig. 5). The spatiotemporal model was able to estimate the total abundance and spawning stock biomass relatively well for all years except year 18.

The comparison between estimated size-specific selectivity from the spatially-aggregated model and true population-level selectivity suggests that the underestimation of abundance is likely due to biased estimates of aggregate selectivity (Fig. 6). The fishing mortality for size class 4 is frequently underestimated. However, the spatiotemporal model was able to estimate the population-level selectivity well (Fig. 6). Although selectivity in the OM is asymptotic (i.e., a logistic curve), selectivity for some years suggests the spatially-aggregated population selection from the spatial OM can be dome-shaped, e.g., years 8 and 16. This could be captured by spatiotemporal model, but not the spatially-aggregated model (Fig. 6).

3.3. Experiment #3: Impact of changing sample sizes on model performance

As the number of sampled locations increases from 50, 100, or 200 per year (Fig 7 for the two species), the convergence rate increases, the average relative error decreases, and the relative bias approaches zero. The convergence rates for northern shrimp are 68%, 72%, and 82% for the data poor, data moderate and data rich scenarios, respectively. The corresponding convergence rates for snow crab case are lower, i.e., 54%, 64%, and 74%. The RMSE appears to decrease as the square-root of annual sample sizes as predicted by maximum likelihood asymptotic theory (RMSE for northern shrimp: 6.1%, 3.9%, and 2.8%; RMSEs for snow crab: 3.8%, 2.8%, and 2.1%). The model performs worse for northern shrimp than for snow crab (Table 2; Fig. 7). Also, the model slightly overestimated the abundance of size classes 4 and 5 for northern shrimp (Fig. 7).

4. DISCUSSION

Managed populations exhibit heterogeneous and complex spatial structure, which is often overlooked in modeling their population dynamics for management and conservation (Crone, 2016; Goethel et al., 2011; Turner et al., 1995). We developed a size-structured spatiotemporal model for estimating fine spatial scale population dynamics and anthropogenic impacts, i.e., fishery dynamics, and used two marine invertebrates with different types of life history to demonstrate our modeling approach. This spatiotemporal model produced unbiased estimates of abundance and fishing mortality spatially and outperformed a spatially-aggregated model when time-varying selectivity caused by spatial heterogeneity in fishing pressure is ignored. To our knowledge, this is the first study to use a simulation experiment to compare the performance of spatiotemporal and spatially-aggregated models that include fishery harvest. Our modeling approach bridges the gap between species distribution and population dynamic models and provides the opportunity to improve natural resource management and conservation by explicitly modeling species' spatiotemporal population and anthropogenic dynamics.

We have demonstrated that our modeling framework can be adapted to populations with different types of life history through straightforward modifications to the population dynamic component, i.e., $g(\mathbf{n}_{s,t})$. Although we presented a size-structured model, it is fairly easy to modify $g(\mathbf{n}_{s,t})$ to an age-structured model. The modeling framework is also flexible enough to accommodate varying degrees of model complexity. In the simplest scenario, the number of size/age classes can be reduced to one and the population dynamic component can be modified to be a delay-difference (Thorson et al., 2015) or biomass dynamic model (Thorson, Jannot, & Somers, 2017). Furthermore, environmental covariates, e.g., sea surface/bottom temperature, salinity, and etc., can be easily incorporated in the model. For example, covariates can be added to Eq. 1 as predictors of density or added in the observation model that relates observations to predicted catch rate, assuming they affect catchability. With additional environmental data, we hypothesize that the model would produce more precise estimates and/or more parameters can be estimated. Therefore, hypotheses such as environmentally-driven recruitment, impact of habitat loss and climate change, and climate change-related distribution shifts can be examined directly within the modeling framework at a fine spatial scale, which may be more useful than approaches that treat model estimates as data for subsequent analysis and rely upon spatially-aggregated data. For example, recruitment and spawning stock biomass estimates from a stock assessment model were used to examine the impacts of environmental variables (Cooper, Rogers,

& Wilderbuer, 2019). Finally, movement can be explicitly incorporated to the model as we did in the OM (Eq. 20). The movement function in the OM accounts for both diffusive and advective movement patterns. However, the simulation scenarios we tested in this study only had diffusive movement. Future research could seek to evaluate different movement patterns, e.g., ontogenic movement, which would also impact the selectivity. It is common for aquatic and terrestrial animals to have movement patterns varying with ontogeny. For example, older birds with more experience are more likely to innovate new migration patterns in response to global change (Teitelbaum et al., 2016).

Integral projection models (IPMs), which predict vital rates from state variables (e.g., size, weight, or age) and covariates (e.g., environment) using regression models, have been increasingly applied to animal and plant populations (e.g., Coulson et al. 2010, Jongejans et al. 2011, Coulson 2012, Merow et al. 2014). These models are considered to have strengths compared to traditional matrix population models (Merow et al., 2014). However, spatial heterogeneity, as one of the most important factors influencing population dynamics, is much less often incorporated into IPMs. Crone (2016) found that spatial heterogeneity increased population growth rates of pasqueflower and suggested that it is important to consider spatial heterogeneity when modeling plant population dynamics. Research efforts have been made to link IPMs with dispersal to model spatial spread (Jongejans et al., 2011). We envision that the similar spatiotemporal modelling approach as we proposed here could be an interesting avenue for future research of IPMs.

Spatially-explicit population dynamics models are increasingly structured using multiple spatial strata (Goethel & Berger, 2017; Goethel et al., 2011). However, this approach requires extensive data to allow a fine spatial scale because (1) each stratum needs sufficient data so that model is tractable, and (2) additional data are often needed to estimate or predefine the connectivity among strata, e.g., movement. However, our modeling framework relies upon totally different structure and assumption, i.e., population density is continuous across the whole area, and estimates the density fields based on geostatistical theory. Therefore, spatial-referenced data from fishery-independent and -dependent survey can be directly used in the model. Our experiment involving different sample sizes shows that the spatiotemporal model can perform well with as few as 100 samples per year. Few studies have been conducted to investigate this approach (Kristensen et al., 2014; Thorson et al., 2015). Kristensen et al. (2014) demonstrates

641 that it is feasible to combine stock assessment and spatiotemporal dynamics. However, they did
642 not include fisheries in their model. Thorson et al. (2015) used a similar modeling approach to
643 estimate a spatially-explicit delay-difference dynamic of a fish population. Our study is an
644 extension of these two and provides a more general modeling framework and rigorous model
645 evaluation. Spatiotemporal models can also be used as operating models for conducting
646 management strategy evaluation (Boyd, Roy, Sibly, Thorpe, & Hyder, 2018), evaluating the
647 performance of stock assessment models, and optimizing sampling design.
648 Spatially-stratified models do not always outperform spatially-implicit/spatially-
649 aggregated models (Szuwalski & Punt, 2015). We suspect that, when there is uncertainty in
650 population spatial structure and movement, incorrect assumption of boundaries between sub-
651 stocks (e.g., mis-specified spatial strata) would lead to poorer estimation. However, the
652 spatiotemporal model of this paper does not rely on spatial strata and accounts for movement
653 implicitly without requiring additional data. Based on our simulation study, we also found that
654 when there is no complex spatial structure, the spatially-aggregated model had similar
655 performance to the spatiotemporal model in estimating spatially-aggregated population and
656 fishery quantities. The comparison scenario we show here represents the situation where a strong
657 and persistent gradient of fishing pressure occurs over space and time. In this case, a spatially-
658 aggregated model could not accurately estimate the population-level fishing mortality. Similar
659 results have been found in Sampson et al. (2011) where an age-structured model was used.
660 Therefore, we envision that our spatiotemporal modeling approach would be especially useful
661 for species that have protected areas or where population pressures vary substantially across
662 space. For instance, spatially-explicit population models have been used to evaluate habitat
663 restoration for cactus wren (*Campylorhynchus brunneicapillus sandiegensis*) (Conlisk, Motheral,
664 Chung, Wisinski, & Endress, 2014). Also, spatiotemporal models can be used to determine the
665 abundance and spatial pattern for endangered species such as amur tiger (*Panthera tigris altaica*)
666 and leopard (*Panthera pardus orientalis*) (Wang et al., 2016) so that effective conservation plans
667 can be developed.
668 We showed that the spatiotemporal model outperforms spatially-aggregated model
669 because of its ability to attribute changes in selectivity to spatial patterns in fishery exploitation.
670 Specifically, the spatially-aggregated model assumed that fishery selectivity was constant over
671 time and space, and this specification was a poor approximation to fishery removals. It is well

known that misspecification of selectivity would lead to biased estimates of population quantities (Linton & Bence, 2011; Stewart & Martell, 2014), and some fishery stock assessments specify time-varying selectivity (Martell & Stewart, 2014). Therefore, spatially-aggregated models that estimate time-varying selectivity may produce less biased results e.g., using time-varying selectivity and accounting for autocorrelation among size and time (Xu, Thorson, Methot, & Taylor, 2018). However, this approach to time-varying selectivity requires estimating a process (aggregate fishery selectivity) that cannot be corroborated through any field sampling, whereas the spatio-temporal model used here approximates the same process by estimating spatial variation in population density, and we argue that the latter is superior because (1) it is more biologically interpretable and (2) could be corroborated by other field sampling.

The convergence rates of the model are not ideal, i.e., 82% and 74% for data rich scenarios (i.e., 200 sampling locations per year) for northern shrimp and snow crab, respectively. This implies that the model requires high-quality and -quantity data. For each iteration, the survey data were re-drawn from the spatial domain. Therefore, the spatial coverage and locations of samples might have an impact on the model convergence rate. As the number of sampling locations decreases, the spatial coverage of samples decreases. This might also be the reason why the convergence rates for snow crab case are lower. The total number of grids used for snow crab simulation is about an order of magnitude greater than that used in northern shrimp. Therefore, with the same number of sampling locations randomly selected from the grids, the samples for northern shrimp are more likely to have a better spatial coverage and be more informative. We also found that the model was much easier to converge when selectivity parameters were fixed.

We conclude that spatially-explicit population models can provide valuable insights into population dynamics and spatial distribution that are not possible with either spatially-aggregated models or species distribution models in isolation, and are useful tools for population ecologists, conservation biologists, and land managers. This advance comes at the expense of greater data requirements. Challenges remain when it comes to application because the model is complex and requires detailed spatially-referenced fishery-dependent and -independent data. Furthermore, it is challenging to define and calculate biological reference points and hence determine associated catch quotas in a spatial context. Spatial harvest strategies can be evaluated using projections (Bosley et al., 2019), which we intend to do using our spatiotemporal model in future work.

Although we argue that when there are survey demographic data available, it can be advantageous to investigate size-structured spatiotemporal models, the spatially-aggregated models can perform similarly as the spatially-explicit models in terms of tracking the whole population.

ACKNOWLEDGEMENTS

We thank K. Kristensen and the developers of Template Model Builder, without which this analysis would not be feasible. J. Cao was supported by NOAA "Stock Assessment and Analytic Methods" (SAAM) grant awarded in 2016 to J. Thorson, A. Punt, and C. Szuwalski. This publication is funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement NA15OAR4320063. We thank R. Methot for helpful comments on an earlier draft. Finally, we would like to thank Dr. Daniel Goethel and two anonymous reviewers whose comments and suggestions helped improve the manuscript.

DATA AVAILABILITY STATEMENT

The data and codes that support the findings of this study are either published or available from the corresponding author upon reasonable request.

REFERENCES

Abbott, J. K., Garber-Yonts, B., & Wilen, J. E. (2010). Employment and Remuneration Effects of IFQs in the Bering Sea/Aleutian Islands Crab Fisheries. *Marine Resource Economics*, *25*(4), 333–354. https://doi.org/10.5950/0738-1360-25.4.333

Adams, L. G., Stephenson, R. O., Dale, B. W., Ahgook, R. T., & Demma, D. J. (2008). Population Dynamics and Harvest Characteristics of Wolves in the Central Brooks Range, Alaska. *Wildlife Monographs*, *170*(1), 1–25. https://doi.org/10.2193/2008-012

Benson, A. J., Cox, S. P., & Cleary, J. S. (2015). Evaluating the conservation risks of aggregate harvest management in a spatially-structured herring fishery. *Fisheries Research*, *167*, 101–113. https://doi.org/10.1016/J.FISHRES.2015.02.003

Berger, A. M., Jones, M. L., Zhao, Y., & Bence, J. R. (2012). Accounting for spatial population structure at scales relevant to life history improves stock assessment: The case for Lake Erie

730 walleye Sander vitreus. *Fisheries Research*, *115–116*, 44–59.
731 https://doi.org/10.1016/j.fishres.2011.11.006

732 Berkeley, S. A., Hixon, M. A., Larson, R. J., & Love, M. S. (2004). Fisheries Sustainability via
733 Protection of Age Structure and Spatial Distribution of Fish Populations. *Fisheries*, *29*(8),
734 23–32. https://doi.org/10.1577/1548-8446(2004)29[23:FSVPOA]2.0.CO;2

735 Beverton, R. J. H., & Holt, S. J. (1957). *On the Dynamics of Exploited Fish Populations*
736 (Investigat). https://doi.org/10.1007/978-94-011-2106-4

737 Bieber, C., & Ruf, T. (2005). Population dynamics in wild boar Sus scrofa : ecology, elasticity of
738 growth rate and implications for the management of pulsed resource consumers. *Journal of*
739 *Applied Ecology*, *42*(6), 1203–1213. https://doi.org/10.1111/j.1365-2664.2005.01094.x

740 Bosley, K. M., Goethel, D. R., Berger, A. M., Deroba, J. J., Fenske, K. H., Hanselman, D. H., …
741 Schueller, A. M. (2019). Overcoming challenges of harvest quota allocation in spatially
742 structured populations. *Fisheries Research*, *220*, 105344.
743 https://doi.org/10.1016/j.fishres.2019.105344

744 Boudreau, S. A., Shackell, N. L., Carson, S., & den Heyer, C. E. (2017). Connectivity,
745 persistence, and loss of high abundance areas of a recovering marine fish population in the
746 Northwest Atlantic Ocean. *Ecology and Evolution*, *7*(22), 9739–9749.
747 https://doi.org/10.1002/ece3.3495

748 Boyd, R., Roy, S., Sibly, R., Thorpe, R., & Hyder, K. (2018). A general approach to
749 incorporating spatial and temporal variation in individual-based models of fish populations
750 with application to Atlantic mackerel. *Ecological Modelling*, *382*, 9–17.
751 https://doi.org/10.1016/J.ECOLMODEL.2018.04.015

752 Cao, J., Chen, Y., & Richards, R. A. (2017a). Improving assessment of *Pandalus* stocks using a
753 seasonal, size-structured assessment model with environmental variables. Part I: Model
754 description and application. *Canadian Journal of Fisheries and Aquatic Sciences*, *74*(3),
755 349–362. https://doi.org/10.1139/cjfas-2016-0020

756 Cao, J., Chen, Y., & Richards, R. A. (2017b). Improving assessment of *Pandalus* stocks using a

757 seasonal, size-structured assessment model with environmental variables. Part II: Model
758 evaluation and simulation. *Canadian Journal of Fisheries and Aquatic Sciences*, *74*(3),
759 363–376. https://doi.org/10.1139/cjfas-2016-0021

760 Conlisk, E., Motheral, S., Chung, R., Wisinski, C., & Endress, B. (2014). Using spatially-explicit
761 population models to evaluate habitat restoration plans for the San Diego cactus wren
762 (Campylorhynchus brunneicapillus sandiegensis). *Biological Conservation*, *175*, 42–51.
763 https://doi.org/10.1016/J.BIOCON.2014.04.010

764 Conroy, M. J., Cohen, Y., James, F. C., Matsinos, Y. G., & Maurer, B. A. (1995). Parameter
765 Estimation, Reliability, and Model Improvement for Spatially Explicit Models of Animal
766 Populations. *Ecological Applications*, *5*(1), 17–19. https://doi.org/10.2307/1942047

767 Cooper, D., Rogers, L. A., & Wilderbuer, T. (2019). Environmentally-driven forecasts of
768 northern rock sole (*Lepidopsetta polyxystra*) recruitment in the eastern Bering Sea.
769 *Fisheries Oceanography*, fog.12458. https://doi.org/10.1111/fog.12458

770 Coulson, T. (2012). Integral projections models, their construction and use in posing hypotheses
771 in ecology. *Oikos*, *121*(9), 1337–1350. https://doi.org/10.1111/j.1600-0706.2012.00035.x

772 Coulson, T., Tuljapurkar, S., & Childs, D. Z. (2010). Using evolutionary demography to link life
773 history theory, quantitative genetics and population ecology. *Journal of Animal Ecology*,
774 *79*(6), 1226–1240. https://doi.org/10.1111/j.1365-2656.2010.01734.x

775 Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. Ver, & Wikle, C. K. (2009). Accounting for
776 uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical
777 modeling. *Ecological Applications*, *19*(3), 553–570. https://doi.org/10.1890/07-0744.1

778 Crone, E. E. (2016). Contrasting effects of spatial heterogeneity and environmental stochasticity
779 on population dynamics of a perennial wildflower. *Journal of Ecology*, *104*(2), 281–291.
780 https://doi.org/10.1111/1365-2745.12500

781 Dunn, A., Rasmussen, S., & Mormede, S. (2014). *Spatial Population Model User Manual*.

782 Dunning, J. B., Stewart, D. J., Danielson, B. J., Noon, B. R., Root, T. L., Lamberson, R. H., &

783 Stevens, E. E. (1995). Spatially Explicit Population Models: Current Forms and Future Uses.
784 *Ecological Applications*, *5*(1), 3–11. https://doi.org/10.2307/1942045

785 Ehrlén, J., & Morris, W. F. (2015). Predicting changes in the distribution and abundance of
786 species under environmental change. *Ecology Letters*, *18*(3), 303–314.
787 https://doi.org/10.1111/ele.12410

788 Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and
789 Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*,
790 *40*(1), 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

791 Fournier, D. A., Hampton, J., & Sibert, J. R. (1998). MULTIFAN-CL: A length-based, age-
792 structured model for fisheries stock assessment, with application to South Pacific albacore,
793 Thunnus alalunga. *Canadian Journal of Fisheries and Aquatic Sciences*, *55*(9), 2105–2116.
794 https://doi.org/10.1139/f98-100

795 Fu, C., & Fanning, L. P. (2004). Spatial Considerations in the Management of Atlantic Cod off
796 Nova Scotia, Canada. *North American Journal of Fisheries Management*, *24*(3), 775–784.
797 https://doi.org/10.1577/m03-134.1

798 Goethel, D. R., & Berger, A. M. (2017). Accounting for spatial complexities in the calculation of
799 biological reference points: effects of misdiagnosing population structure for stock status
800 indicators. *Canadian Journal of Fisheries and Aquatic Sciences*, *74*(11), 1878–1894.
801 https://doi.org/10.1139/cjfas-2016-0290

802 Goethel, D. R., Legault, C. M., & Cadrin, S. X. (2014). Demonstration of a spatially explicit,
803 tag-integrated stock assessment model with application to three interconnected stocks of
804 yellowtail flounder off of New England. *ICES Journal of Marine Science*, *72*(1), 164–177.
805 https://doi.org/10.1093/icesjms/fsu014

806 Goethel, D. R., Legault, C. M., & Cadrin, S. X. (2015). Testing the performance of a spatially
807 explicit tag-integrated stock assessment model of yellowtail flounder (*Limanda ferruginea*)
808 through simulation analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, *72*(4),
809 582–601. https://doi.org/10.1139/cjfas-2014-0244

Goethel, D. R., Quinn, T. J., & Cadrin, S. X. (2011). Incorporating Spatial Structure in Stock Assessment: Movement Modeling in Marine Fish Population Dynamics. *Reviews in Fisheries Science*, *19*(2), 119–136. https://doi.org/10.1080/10641262.2011.557451

Guan, W., Cao, J., Chen, Y., & Cieri, M. (2013). Impacts of population and fishery spatial structures on fishery stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, *70*(8), 1178–1189. https://doi.org/10.1139/cjfas-2012-0364

Guisan, A., Edwards, T. C., & Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, *157*(2–3), 89–100. https://doi.org/10.1016/S0304-3800(02)00204-1

Hulson, P. J. F., Miller, S. E., Ianelli, J. N., & Quinn, T. J. (2011). Including mark-recapture data into a spatial agestructured model: Walleye Pollock (Theragra chalcogramma) in the eastern Bering Sea. *Canadian Journal of Fisheries and Aquatic Sciences*, *68*(9), 1625–1634. https://doi.org/10.1139/f2011-060

Hurtado-Ferro, F., Punt, A. E., & Hill, K. T. (2014). Use of multiple selectivity patterns as a proxy for spatial structure. *Fisheries Research*, *158*, 102–115. https://doi.org/10.1016/j.fishres.2013.10.001

Jongejans, E., Shea, K., Skarpaas, O., Kelly, D., & Ellner, S. P. (2011). Importance of individual and environmental variation for invasive species spread: a spatial integral projection model. *Ecology*, *92*(1), 86–97. https://doi.org/10.1890/09-2226.1

Jongejans, E., Skarpaas, O., & Shea, K. (2008). Dispersal, demography and spatial population models for conservation and control management. *Perspectives in Plant Ecology, Evolution and Systematics*, *9*(3–4), 153–170. https://doi.org/10.1016/J.PPEES.2007.09.005

Kerr, L. A., Cadrin, S. X., & Secor, D. H. (2010). The role of spatial dynamics in the stability, resilience, and productivity of an estuarine fish population. *Ecological Applications*, *20*(2), 497–507. https://doi.org/10.1890/08-1382.1

Krebs, C. J. (1972). *Ecology : the experimental analysis of distribution and abundance*. New York, NY: HarperCollins College Publishers.

837 Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., & Bell, B. (2015). TMB: automatic
838 differentiation and Laplace approximation. *ArXiv Preprint ArXiv:1509.00660*.

839 Kristensen, K., Thygesen, U. H., Andersen, K. H., & Beyer, J. E. (2014). Estimating spatio-
840 temporal dynamics of size-structured populations. *Canadian Journal of Fisheries and*
841 *Aquatic Sciences*, *71*(2), 326–336. https://doi.org/10.1139/cjfas-2013-0151

842 Kuo, T.-C., Mandal, S., Yamauchi, A., & Hsieh, C. (2015). Life history traits and exploitation
843 affect the spatial mean-variance relationship in fish abundance. *Ecology*, *97*(5), 15-1270.1.
844 https://doi.org/10.1890/15-1270.1

845 Lee, H.-H., Piner, K. R., Maunder, M. N., Taylor, I. G., & Methot, R. D. (2017). Evaluation of
846 alternative modelling approaches to account for spatial effects due to age-based movement.
847 *Canadian Journal of Fisheries and Aquatic Sciences*, *74*(11), 1832–1844.
848 https://doi.org/10.1139/cjfas-2016-0294

849 Linton, B. C., & Bence, J. R. (2011). Catch-at-age assessment in the face of time-varying
850 selectivity. *ICES Journal of Marine Science*, *68*(3), 611–625.
851 https://doi.org/10.1093/icesjms/fsq173

852 Martell, S., & Stewart, I. (2014). Towards defining good practices for modeling time-varying
853 selectivity. *Fisheries Research*, *158*, 84–95.
854 https://doi.org/10.1016/J.FISHRES.2013.11.001

855 Maunder, M. N., & Piner, K. R. (2015). Contemporary fisheries stock assessment: many issues
856 still remain. *ICES Journal of Marine Science*, *72*(1), 7–18.
857 https://doi.org/10.1093/icesjms/fsu015

858 Merow, C., Dahlgren, J. P., Metcalf, C. J. E., Childs, D. Z., Evans, M. E. K., Jongejans, E., …
859 McMahon, S. M. (2014). Advancing population ecology with integral projection models: a
860 practical guide. *Methods in Ecology and Evolution*, *5*(2), 99–110.
861 https://doi.org/10.1111/2041-210X.12146

862 Mueter, F. J., & Litzow, M. A. (2008). Sea ice retreat alters the biogeography of the Bering Sea
863 continental shelf. *Ecological Applications*, *18*(2), 309–320. https://doi.org/10.1890/07-

0564.1

Parada, C., Armstrong, D. A., Ernst, B., Hinckley, S., & Orensanz, J. M. (2010). Spatial dynamics of snow crab (Chionoecetes opilio) in the eastern Bering Sea—putting together the pieces of the puzzle. *Bulletin of Marine Science*, *86*(2), 413–437.

Punt, A. E. (2019). Spatial stock assessment methods: A viewpoint on current issues and assumptions. *Fisheries Research*, *213*, 132–143. https://doi.org/10.1016/j.fishres.2019.01.014

Punt, A. E., Haddon, M., Little, L. R., & Tuck, G. N. (2016). Can a spatially-structured stock assessment address uncertainty due to closed areas? A case study based on pink ling in Australia. *Fisheries Research*, *175*, 10–23. https://doi.org/10.1016/J.FISHRES.2015.11.008

Punt, A. E., Haddon, M., & Tuck, G. N. (2015). Which assessment configurations perform best in the face of spatial heterogeneity in fishing mortality, growth and recruitment? A case study based on pink ling in Australia. *Fisheries Research*, *168*, 85–99. https://doi.org/10.1016/j.fishres.2015.04.002

Quinn, T. J., Deriso, R. B., & Neal, P. R. (1990). Migratory catch-age analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, *47*(12), 2315–2327. https://doi.org/10.1139/f90-258

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sampson, D. B., & Scott, R. D. (2011). A spatial model for fishery age-selection at the population level. *Canadian Journal of Fisheries and Aquatic Sciences*, *68*(6), 1077–1086. https://doi.org/10.1139/f2011-044

Sampson, D. B., & Scott, R. D. (2012). An exploration of the shapes and stability of population-selection curves. *Fish and Fisheries*, *13*(1), 89–104. https://doi.org/10.1111/j.1467-2979.2011.00417.x

Sampson, D. B., Scott, R. D., & Quinn, T. (2011). A spatial model for fishery age-selection at

the population level. *Canadian Journal of Fisheries and Aquatic Sciences*, *68*(6), 1077–1086. https://doi.org/10.1139/f2011-044

Stewart, I. J., & Martell, S. J. D. (2014). A historical review of selectivity approaches and retrospective patterns in the Pacific halibut stock assessment. *Fisheries Research*, *158*, 40–49. https://doi.org/10.1016/J.FISHRES.2013.09.012

Szuwalski, C. S., & Punt, A. E. (2015). Can an aggregate assessment reflect the dynamics of a spatially structured stock? Snow crab in the eastern Bering Sea as a case study. *Fisheries Research*, *164*, 135–142. https://doi.org/10.1016/J.FISHRES.2014.10.020

Teitelbaum, C. S., Converse, S. J., Fagan, W. F., Böhning-Gaese, K., O'Hara, R. B., Lacy, A. E., & Mueller, T. (2016). Experience drives innovation of new migration patterns of whooping cranes in response to global change. *Nature Communications*, *7*, 12793. https://doi.org/10.1038/ncomms12793

Thorson, J. T. (2017). Three problems with the conventional delta-model for biomass sampling data, and a computationally efficient alternative. *Canadian Journal of Fisheries and Aquatic Sciences*, (ja).

Thorson, J. T., & Haltuch, M. A. (2018). Spatio-temporal analysis of compositional data: increased precision and improved workflow using model-based inputs to stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, cjfas-2018-0015. https://doi.org/10.1139/cjfas-2018-0015

Thorson, J. T., Ianelli, J. N., Munch, S. B., Ono, K., & Spencer, P. D. (2015). Spatial delay-difference models for estimating spatiotemporal variation in juvenile production and population abundance. *Canadian Journal of Fisheries and Aquatic Sciences*, *72*(12), 1897–1915. https://doi.org/10.1139/cjfas-2014-0543

Thorson, J. T., Jannot, J., & Somers, K. (2017). Using spatio-temporal models of population growth and movement to monitor overlap between human impacts and fish populations. *Journal of Applied Ecology*, *54*(2), 577–587. https://doi.org/10.1111/1365-2664.12664

Thorson, J. T., Munch, S. B., & Swain, D. P. (2017). Estimating partial regulation in

spatiotemporal models of community dynamics. *Ecology*, *98*(5), 1277–1289.

Turner, M. G., Arthaud, G. J., Engstrom, R. T., Hejl, S. J., Liu, J., Loeb, S., & McKelvey, K. (1995). Usefulness of Spatially Explicit Population Models in Land Management. *Ecological Applications*, *5*(1), 12–16. https://doi.org/10.2307/1942046

Vincent, M. T., Brenden, T. O., & Bence, J. R. (2016). Simulation testing the robustness of a multiregion, tag-integrated assessment model that exhibits natal homing and estimates natural mortality and reporting rate. *Canadian Journal of Fisheries and Aquatic Sciences*, *74*(11), 1930–1949. https://doi.org/10.1139/cjfas-2016-0297

Wang, T., Feng, L., Mou, P., Wu, J., Smith, J. L. D., Xiao, W., … Ge, J. (2016). Amur tigers and leopards returning to China: direct evidence and a landscape conservation plan. *Landscape Ecology*, *31*(3), 491–503. https://doi.org/10.1007/s10980-015-0278-1

Xu, H., Thorson, J. T., Methot, R. D., & Taylor, I. G. (2018). A new semi-parametric method for autocorrelated age- and time-varying selectivity in age-structured assessment models. *Canadian Journal of Fisheries and Aquatic Sciences*, cjfas-2017-0446. https://doi.org/10.1139/cjfas-2017-0446

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

Table 1. List of indices used in model descriptions, data used during parameter estimation and simulation, and all parameters (the type of each parameter is listed as estimated (“fixed”, “random”), or calculated from estimated parameters (“derived quantity”)).

Name	Symbol	Type
Sample	i	Index
Location	s	Index
Year	t	Index
Maximum time step	T	Index

Size bin	l	Index
Immaturity	ω	index
Maturity	λ	index
Knot	j	Index
Number of years	τ	Index
Data of *i*th sample at location s and time t	$x(s_i,t_i)$	Data
Area swept for sample i	a_i	Data
Total area associated with knot j	A_j	Data
Maturity at size	$\mathbf{w}$	Data
Sex ratio of the recruits (male)	p^{male}	Data
Growth transition matrix	$\mathbf{G}$	Data
Natural mortality at size	$\mathbf{m}$	Data
Movement matrix	$\mathbf{M}$	Data
Biomass per group of individuals	$\delta(s_i,t_i)$	Data/derived quantity
Biomass	b	Data/derived quantity
Number of animals per area	n	Data/derived quantity
Encounter probability	p	Derived quantity
Recruitment at size for location *s* and year *t*	$\mathbf{r}_{s,t}$	Derived quantity
Selectivity at size	$\mathbf{v}$	Derived quantity
Catch at size for location *s* and year *t*	$\mathbf{c}_{s,t}$	Derived quantity
Aggregate selectivity for size class *l* and year *t*	$s_{l,t}$	Derived quantity
Variance of positive catch rate	σ_D^2	Parameter (fixed)
Initial abundance at size	$\boldsymbol{\varphi}$	Parameter (fixed)
Variance of fishing mortality	σ_f^2	Parameter (fixed)
Logistic selectivity parameter	θ	Parameter (fixed)
Logistic selectivity parameter	l_{50}	Parameter (fixed)
Average recruitment of year *t*	r_t	Parameter (fixed)
Pairwise covariance between any two size classes	$\boldsymbol{\Theta}_{\boldsymbol{L}}$	Parameter (fixed)
Decorrelation distance	κ	Parameter (fixed)
Matérn smoothness parameter	ν	Parameter (fixed)

Process error for year *t*	$\mathbf{E}_t$	Parameter (random)
Observation-level random effect	η_i	Parameter (random)
Fully-selected fishing mortality at location *s*	$\mathbf{f}_s$	Parameter (random)
Abundance at size over time	$\mathbf{N}$	Parameter (random)

939

940

941 Table 2. The average Root-Mean-Square Error (RMSE) and average Relative Bias (RB) across
942 all years of estimated aggregate abundance-at-size for northern shrimp and snow crab in data
943 poor, moderate and rich scenarios.

944

		Northern shrimp		Snow crab	
size class	sample size	RMSE (%)	RB (%)	RMSE (%)	RB (%)
1	50	6.13	-0.9077	4.08	-1.0498
1	100	3.95	-0.6968	3.21	-1.2290
1	200	2.67	-0.5925	2.24	-0.7661
2	50	5.87	-0.1665	4.02	-0.7559
2	100	3.65	-0.3300	2.94	-0.5656
2	200	2.58	-0.2433	2.17	-0.5143
3	50	6.13	0.3595	3.51	-0.3060
3	100	3.77	0.0082	2.48	-0.2885
3	200	2.56	-0.0809	2.02	-0.2478
4	50	6.31	3.0897	3.61	0.5006
4	100	3.84	1.6495	2.59	0.1132
4	200	2.61	0.8710	1.92	0.0059
5	50	6.19	2.4945	3.58	0.4051
5	100	4.26	1.4511	2.53	0.1206
5	200	3.40	0.9986	2.00	0.0917

945

946

947 **FIGURES**

948

949 Figure 1. Comparison of simulated and estimated distribution of size classes 1, 3 and 5 in
950 selected years using data without measurement error and movement (experiment 1).
951

952 Figure 2. Comparison of simulated (red line) and estimated (black dot) spatially-aggregated total
953 abundance (a) and total removals (b) by size class over time for the scenario with no
954 measurement error nor movement (experiment 1).
955

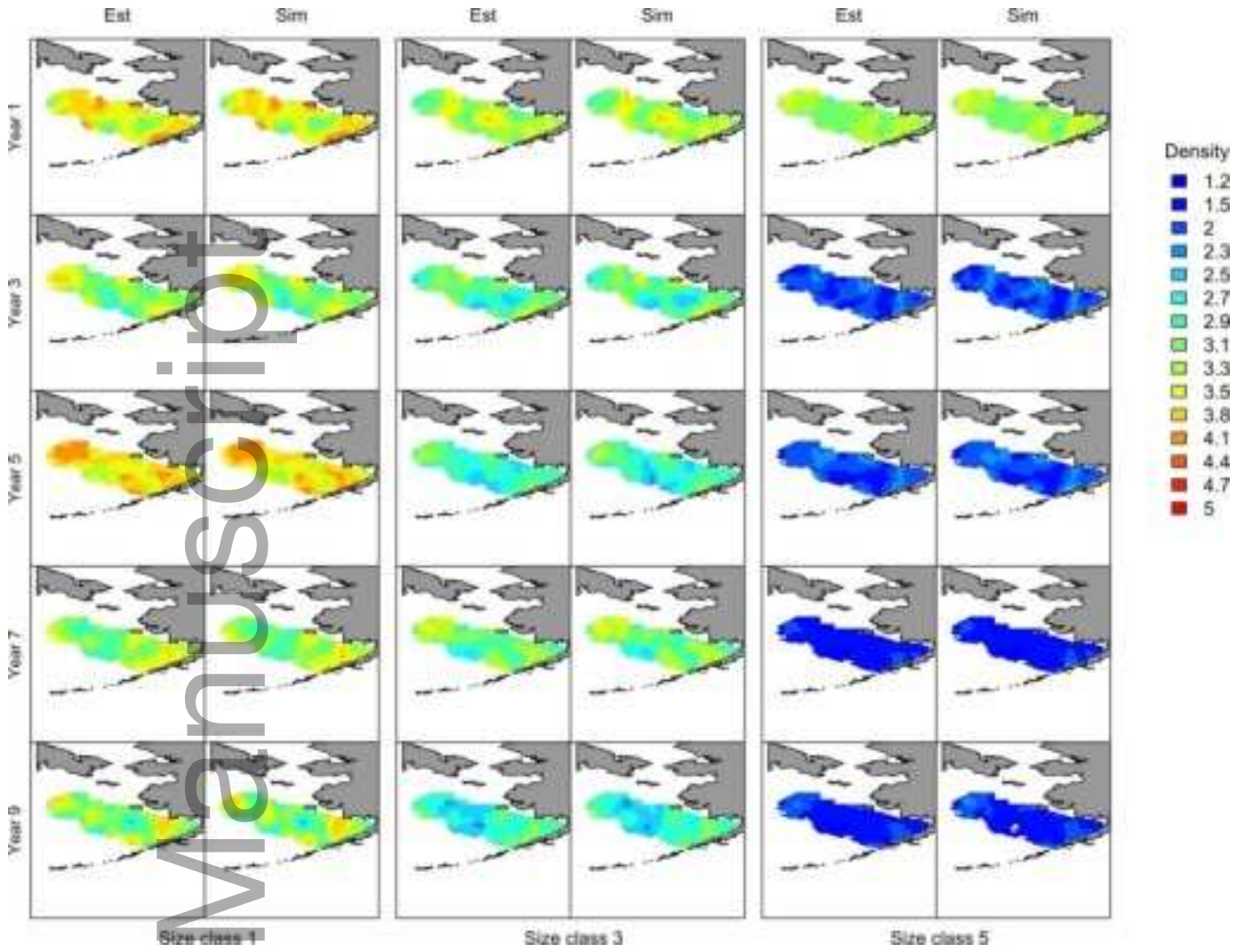
956 Figure 3. Comparison of simulated (red line) and estimated (black dot) spatially-aggregated total
957 abundance (a) and total removals (b) by size class over time for the stochastic data scenario with
958 movement (a randomly selected replicate) (experiment 1). 95% confidence intervals ($\pm$ 1.96∗SE,
959 where SE is the estimated standard error) are shown by error bars. Standard deviations used
960 when generating data are 0.5 for both survey and fishery catches.
961

962 Figure 4. The relative errors (percentages) of aggregate abundance-at-size estimated from
963 spatially-aggregated and spatiotemporal models (experiment 2). Median Root-mean-square error
964 (RMSE) and relative bias (RB) in percentage across years are listed in each panel for comparison.
965

966 Figure 5. The relative errors (percentages) of aggregate total abundance and spawning stock
967 biomass from the spatially-aggregated and spatiotemporal models (experiment 2). Median Root-
968 mean-square error (RMSE) and relative bias (RB) in percentage across years are listed in each
969 panel for comparison.
970

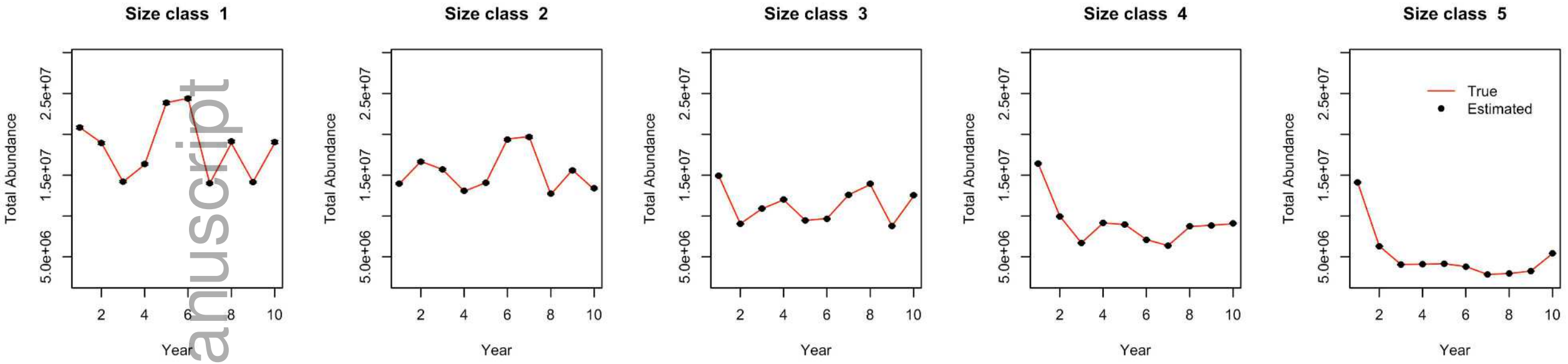
971 Figure 6. The true population-level selectivity at size $s_{l,t}$ (defined in Eq. 21, black dotted line)
972 and 95 percentiles of the estimated selectivity at size from the spatially-aggregated model (red)
973 and spatiotemporal model (yellow) over 50 replicates (experiment 2).
974

975 Figure 7. Boxplot of Root-Mean-Square Error (RMSE) and Relative Bias (RB) of abundance-at-
976 size across replicates for data poor, moderate and rich scenarios, for northern shrimp and snow
977 crab (experiment 3).

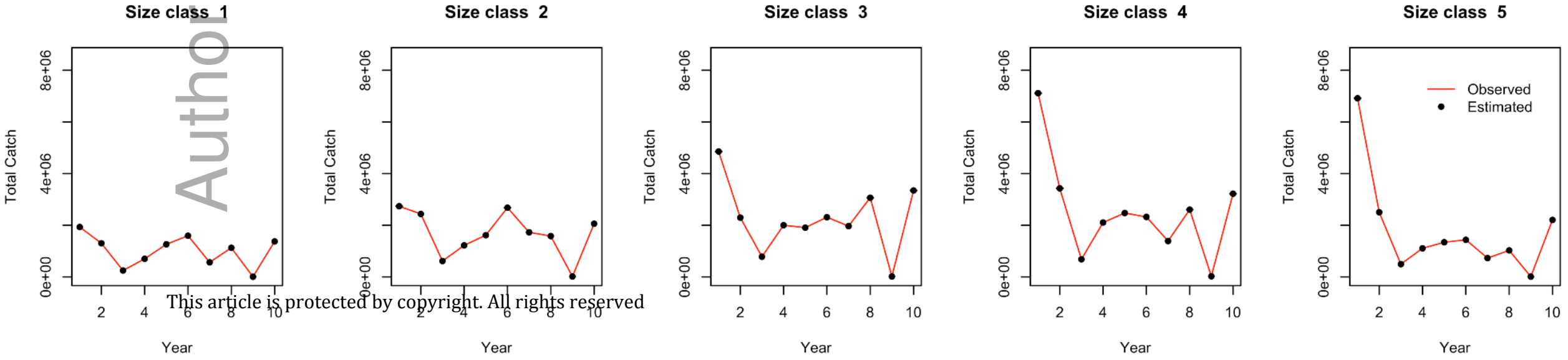


faf_12433_f1.png

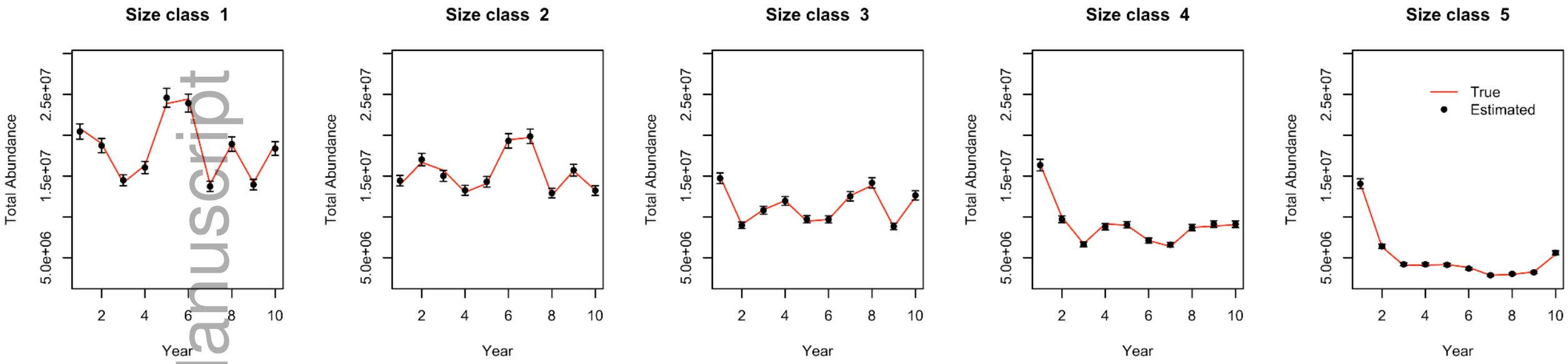
(a)



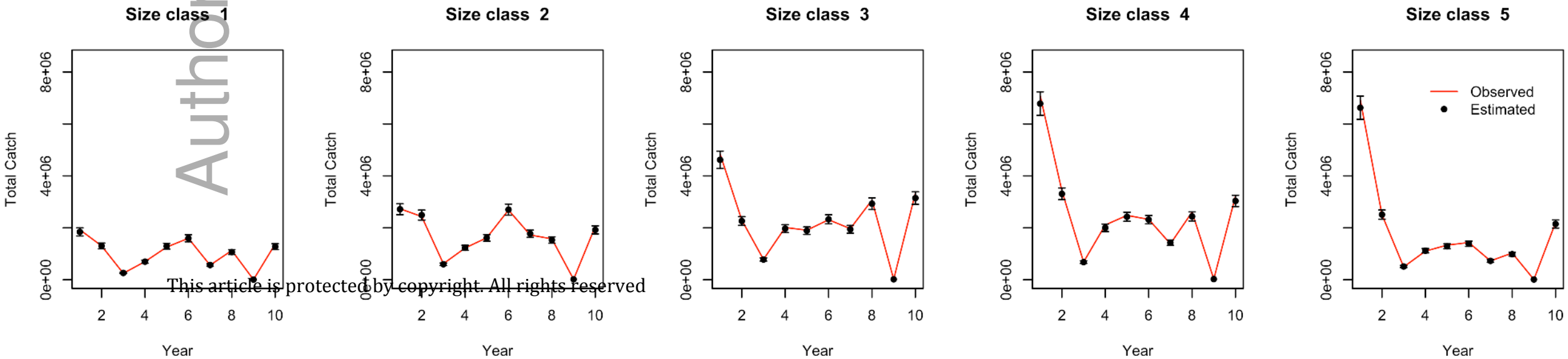
(b)

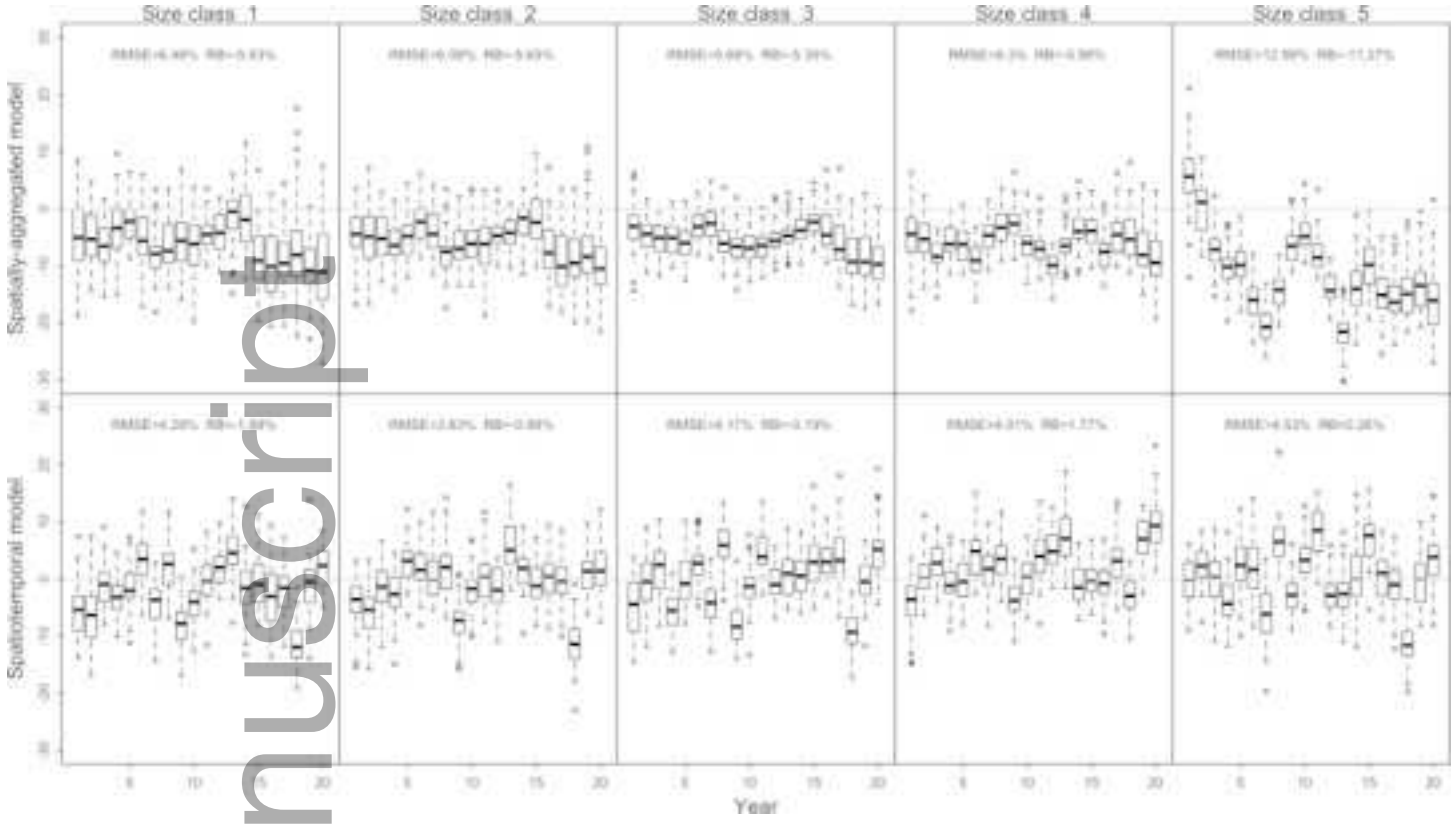


(a)

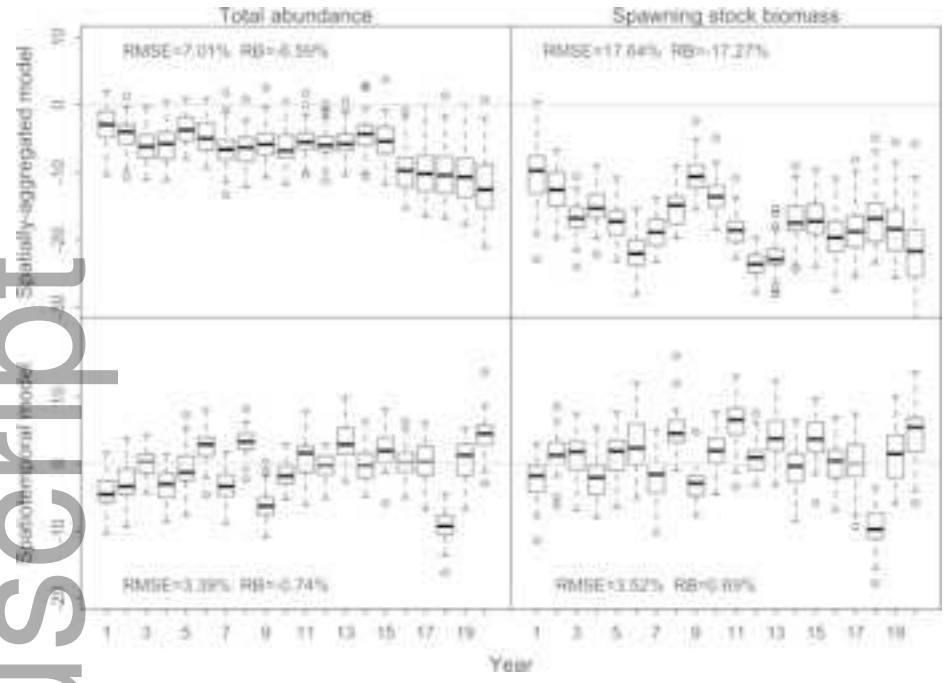


(b)

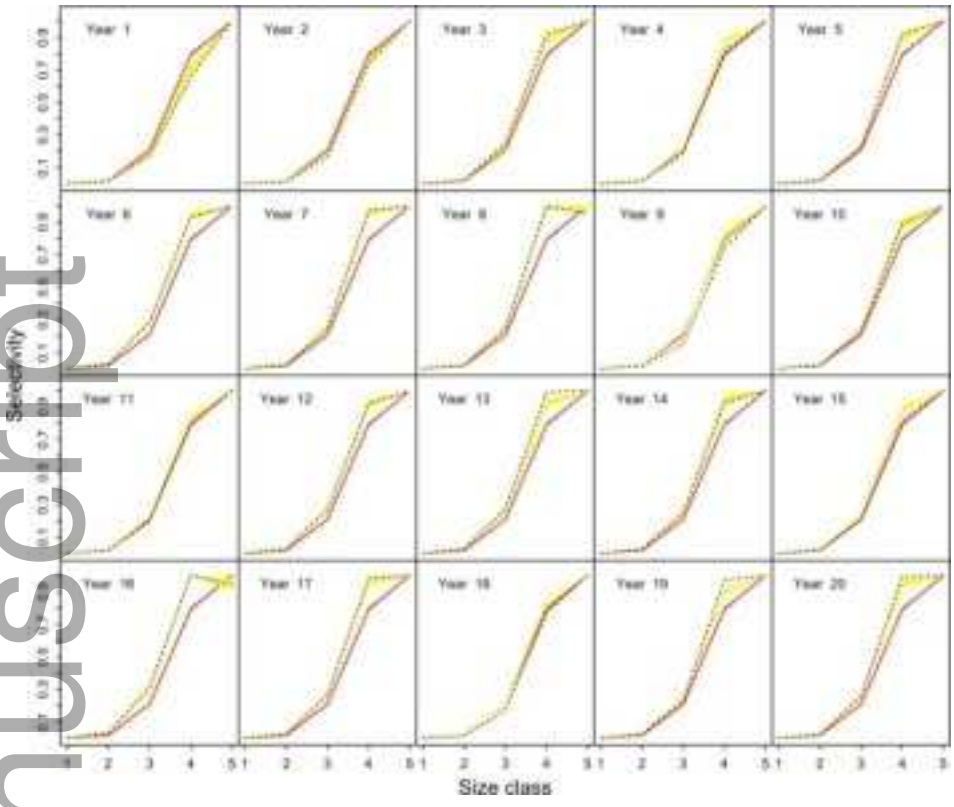




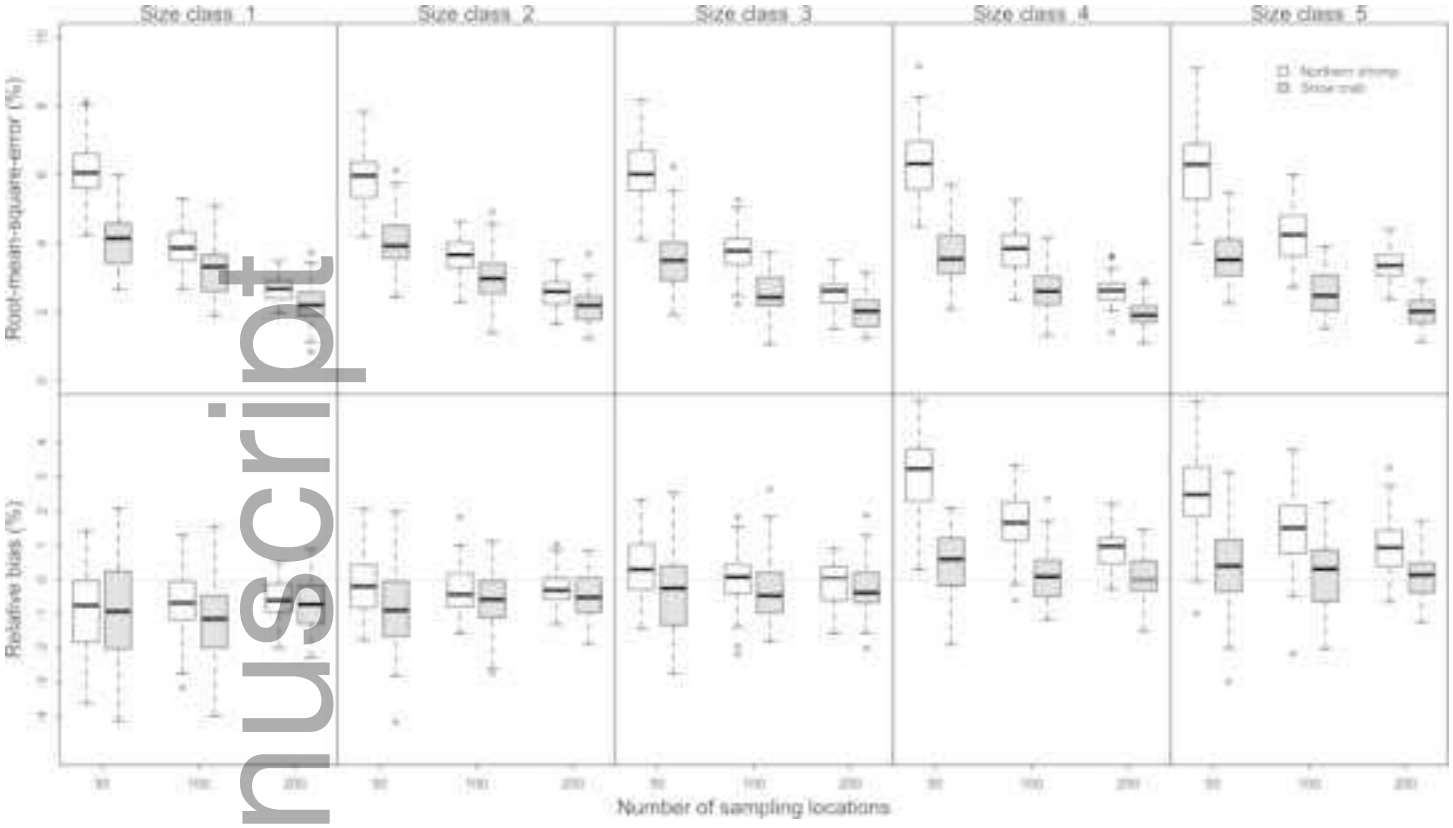
faf_12433_f4.png



faf_12433_f5.png



faf_12433_f6.png



faf_12433_f7.png