MRS. KATELYN BOSLEY (Orcid ID : 0000-0002-5739-6793)
DR. DANIEL GOETHEL (Orcid ID : 0000-0003-0066-431X)

Article type : Original Article

Running Head: Population structure misspecification

Finding the perfect mismatch: Evaluating misspecification of population structure within spatially explicit integrated population models

Katelyn M. Bosley[1,2*], Amy M. Schueller[3], Daniel R. Goethel[4,5], Dana H. Hanselman[5], Kari H. Fenske[5], Aaron M. Berger[2], Jonathan J. Deroba[6], Brian J. Langseth[7,8]

[1]Washington Department of Fish and Wildlife, 375 Hudson St., Port Townsend, WA 98368, USA
[2]Northwest Fisheries Science Center, Fisheries Resource and Monitoring Division, NMFS-NOAA, 2032 S.E. OSU Drive, Newport, OR 97365, USA
[3]Southeast Fisheries Science Center, Beaufort Laboratory, NMFS-NOAA, 101 Pivers Island Road, Beaufort, NC 28516, USA
[4]Southeast Fisheries Science Center, NMFS-NOAA, 75 Virginia Beach Drive, Miami, FL 33133, USA
[5]Alaska Fisheries Science Center, Auke Bay Laboratories, NMFS-NOAA, 17109 Point Lena Loop Road, Juneau, AK 99801, USA
[6]Northeast Fisheries Science Center, NMFS-NOAA, 166 Water St., Woods Hole MA, 02543, USA
[7]Pacific Islands Fisheries Science Center, NMFS-NOAA, 1845 Wasp Blvd., Bldg. 176, Honolulu HI, 96818, USA

[8]Northwest Fisheries Science Center, NMFS-NOAA, 2725 Montlake Blvd E, Seattle, WA, 98112, USA

*Corresponding author

Katelyn Bosley

Current Address: Washington Department of Fish and Wildlife

375 Hudson St., Port Townsend, WA 98368

Email: katelyn.bosley@dfw.wa.gov

Open Research Statement: No empirical data were collected for this study. Novel model code for all simulations is available at https://github.com/KateBoz/Spatial_IPM.

Abstract

Spatially stratified integrated population models (IPMs) can account for fine-scale demographic processes and support spatial management for complex, heterogeneous populations. Although spatial IPMs may provide a more realistic representation of true population dynamics, few studies have evaluated the consequences associated with incorrect assumptions regarding population structure and connectivity. We utilized a simulation-estimation framework to explore how mismatches between the true population structure (i.e., uniform, single population with spatial heterogeneity, or metapopulation) and various parametrizations of an IPM (i.e., panmictic, fleets-as-areas, or a spatially explicit, tag-integrated model) impacted resultant fish population estimates. When population structure was incorrectly specified in the IPM, parameter estimates were generally unbiased at the system level, but were often biased for sub-areas. Correctly specifying population structure in spatial IPMs led to strong performance, while incorrectly specified spatial IPMs performed adequately (and better than spatially aggregated counterparts). Allowing for flexible parametrization of movement rates (e.g., estimating age-varying values) was more important than correctly identifying the population structure, and incorporation of tag-recapture data helped movement estimation. Our results elucidate how incorrect population structure assumptions can influence the estimation of key parameters of spatial IPMs, while indicating that, even if incorrectly specified, spatial IPMs can adequately support spatial management decisions.

Keywords: Alaskan sablefish, fisheries management, movement dynamics, spatial integrated population model, stock assessment, tag-recapture

Table of Contents

1 Introduction

Spatially heterogeneous environments influence almost every aspect of an organism's behavior, ultimately influencing the resultant population dynamics and food-web interactions (Nathan et al., 2008). The impact of spatially dynamic landscapes results in complex population level responses including the formation of metapopulation structure, source-sink dynamics,

predator-prey interactions, heterogeneous life history parameters, and speciation (Keymer et al., 2000; Aguilée et al., 2011; Trainor et al., 2014; Northfield et al., 2017). The discipline of landscape ecology coalesced to explicitly account for spatial heterogeneity in population structure (Pickett and Cadenasso, 1995; Hidalgo et al., 2016), which has resulted in a variety of population-scale spatially explicit ecological models including patch, diffusion, island, species distribution, and continuum models (Hastings, 1990; Kareiva, 1990; Elith and Leathwick, 2009; Trainor et al., 2014; DeAngelis and Yurek, 2017; Northfield et al., 2017). Concomitantly, there has been widespread acknowledgement that accounting for and protecting spatial population structure is critical to maintaining resilient populations, especially those that are directly harvested (e.g., marine fisheries; Ciannelli et al., 2013; Allen and Singh, 2016; Fraser et al., 2018).

In the marine realm, spatial population structure and biocomplexity, as well as misspecification of that complexity, has implications for marine spatial planning and creation of marine protected areas (McGilliard et al., 2015), determination of appropriate management boundary definitions (Berger et al., 2020), and establishment of population status determination criteria (Smedbol and Stephenson, 2001; Cianneli et al., 2013; Goethel and Berger, 2017). For instance, ignoring spatial differences in North Sea Atlantic cod (*Gadus morhua, Gadidae*) populations has been proposed as a potential factor resulting in population collapses (Hutchinson, 2008). Similarly, simulations have shown that explicit spatial management of individual spawning populations was necessary to avoid localized depletion of more vulnerable population components of Atlantic cod off Nova Scotia (Fu and Fanning, 2004) and of small yellow croakers (*Larimichthys polyactis*, Sciaenidae*)* in China (Ying et al., 2011). Kerr et al. (2014) found that the Atlantic cod populations located off the northeastern United States appeared more robust to fishing pressure than when current management boundaries were used to assess the stock rather than the correct biological stock delineations, which could lead to overfishing. Spawning stock biomass and fishing mortality rate were also biased for Atlantic herring (*Clupea harengus,* Clupeidae) when management boundaries were used to assess population status rather than biological boundaries (Guan et al., 2013). Incorrect spatial delineations can also lead to biased regional stock productivity, which has been shown for sardines (*Sardinops sagax,* Clupeidae*)* off South Africa (de Moor and Butterworth, 2015). In the terrestrial realm, spatial population structure and biocomplexity has implications for persistence

given increased habitat fragmentation, predator-prey interactions, and disease transmission (Gu et al., 2002; Trainor et al., 2014; White et al., 2018). For instance, landscape dynamics such as disturbances and successional changes reduced the viability of sharp-tailed grouse (*Tympanuchus phasianellus,* Phasianidae) populations when modeled with population demographics (Akçakaya et al., 2004). Additionally, estimates of survival and reproduction in spatial capture-recapture models for grizzly bears (*Ursus arctos,* Ursidae) were less biased than in non-spatial capture-recapture models (Whittington and Sawaya, 2015).

Incorporating spatial complexity into population dynamics models that directly support management of harvested species remains difficult, because these models must directly estimate population status from limited and uncertain observed data (Struve et al., 2010; Berger et al., 2017; Ogburn et al., 2017). Models of wildlife resource utilization commonly maintain assumptions that the modeled population unit is homogeneously distributed and harvested across the spatial domain, while no immigration or emigration is assumed to occur (Goethel et al., 2011; Chandler and Clark, 2014). In most instances these assumptions lead to model misspecification, because most species demonstrate complex spatial population structure, patchy distributions, connectivity among population or habitat components, spatial variation in life history characteristics, and unequal harvesting across a species' range (Kerr et al., 2017; Zipkin and Saunders, 2018; Punt, 2019b). In the fisheries literature, a common method to implicitly address spatial dynamics in population estimation models is to use the fleets-as-areas (FAA) approach (Punt, 2019b). FAA models assume a single homogenous population unit, but with multiple harvest units (e.g., fishery fleets) that differentially cull segments of populations (i.e., through different size or age selection of individuals), which act as proxies for the spatial structure of the population (Cope and Punt, 2011; Waterhouse et al., 2014). Although FAA models have been shown to outperform naïve spatially explicit models (e.g., Lee et al., 2017), they generally perform no better than spatially aggregated models when complex spatial structure exists (Punt et al., 2015, 2016, 2017; Punt, 2019b).

Application of spatially explicit estimation models often improve estimates of population productivity by simultaneously assessing individual spawning components, as well as connectivity dynamics among them instead of aggregating data and parameter estimates across multiple reproductive units (Chandler and Clark, 2014; Zipkin and Saunders, 2018; Berger et al., 2017; Punt, 2019b). Over the last two decades, the increasing application of integrated

154 population models (IPMs) has allowed incorporation of new and novel data streams (e.g., bio-
155 logging and fine-scale demographic information), thereby increasing the quantity and spatial
156 resolution of data inputs (Maunder and Punt, 2013; Zipkin and Saunders, 2018). IPMs are
157 estimation models that can incorporate spatially and structurally diverse data sets into a unified
158 framework by utilizing a single, combined objective function (Maunder and Punt, 2013). These
159 models have the flexibility to include an array of data sources and multiscalar population
160 processes, which has led to increased implementation of spatially explicit IPMs for both
161 terrestrial and fisheries applications (Chandler and Clark, 2014; Berger et al., 2017; Zipkin and
162 Saunders, 2018). By explicitly modeling spatial dynamics (e.g., spatial variation in demographics
163 and connectivity among spawning populations), spatially explicit IPMs can match the local scale
164 of each data set while identifying regional scale temporal changes in species distributions
165 (Goethel et al., 2021).
166 Although spatial IPMs can better account for biocomplexity compared to spatially
167 aggregated, closed population models, there is a limit to the types of spatial population structure,
168 number of population components, and complexity of connectivity dynamics that can be
169 modeled (Punt, 2019b; Cadrin, 2020). Goethel et al. (2011) present a generalized spatially
170 explicit IPM that can estimate movement rates and account for the three primary types of spatial
171 population structure observed in marine populations: spatial heterogeneity, metapopulation
172 structure, and natal homing. Despite more fine-scale dynamics being likely in many species (e.g.,
173 contingent structure; Petitgas et al., 2010), the parsimony-complexity tradeoff, as well as
174 limitations to the spatial scale of fisheries data, limit the number of population components and
175 spatial areas that can be modeled in an IPM (Punt et al., 2018; Goethel et al., 2021; although see
176 Cao et al., 2020, for a spatiotemporal approach that uses spatial autocorrelation to model
177 population dynamics at an extremely fine-scale). It has been widely demonstrated that assuming
178 homogenous populations or applying spatially aggregated IPMs may be detrimental to achieving
179 sustainable management of marine populations (e.g., Ying et al., 2011; McGilliard et al., 2015;
180 Goethel et al., 2021). Yet, aside from assuming population homogeneity when metapopulation or
181 natal homing structure is present (e.g., Ying et al., 2011; Li et al., 2015, 2018) few studies have
182 explored the management consequences of misdiagnosing the form of the underlying spatial
183 population structure in spatially explicit IPMs across the array of common population structures
184 observed in marine species.

We developed and applied a spatially explicit simulation-estimation framework to explore how misdiagnosis of spatial population structure in marine fish populations can influence estimates of population status when conducting IPMs. Our primary objectives sought to 1) determine the extent of bias in estimates of population status across an array of spatially aggregated, spatially implicit, and spatially explicit IPMs when misspecification of the true population structure exists; 2) establish which spatial IPM specifications are most robust to complex underlying spatial processes; and 3) to explore if more complex parameterizations of movement can help overcome misdiagnosed spatial structure. To improve the realism of the simulation experiment we emulated the spatial dynamics of Alaskan sablefish (*Anoplopoma fimbria,* Anoplopomatidae), which undertake long-distance ontogenetic migrations along the Alaskan coast (Hanselman et al., 2015), have complex management boundaries, and have undergone extensive tag-recapture experiments for over thirty years. The results of this study provide insight into the performance of spatially explicit IPMs when the underlying spatial population structure is poorly understood. This work also provides guidance on parameterization of spatially explicit IPMs for producing robust estimates of population status.

2 Methods

2.1 Overview

A simulation-estimation framework was developed to evaluate the performance of IPMs utilizing a range of assumptions regarding the underlying spatial population structure and movement dynamics. The operating model (OM), representing the true dynamics of the system, was conditioned using parameters that emulate the dynamics of Alaskan sablefish (hereafter referred to as sablefish). Sablefish are relatively long-lived, highly mobile, and inhabit three management areas off the Alaskan coast (See Supplemental Material Fig. B1). Three parametrizations of the OM were developed to represent varying degrees of spatial complexity: 1) a single homogeneously distributed population in all three areas (*Uniform*; akin to a panmictic population), 2) a single heterogeneously distributed population in all three areas (*Spatial Heterogeneous, SH*), and 3) a metapopulation with a different subpopulation in each area (*Metapopulation*). The two spatial OMs included complex time- and age-varying movement patterns among population units among areas. These OMs were used to generate simulated pseudo-data that were fit within each of four separate IPMs, which varied in spatial complexity

(i.e., panmictic, FAA, spatially heterogeneous, and metapopulation). Spatially heterogeneous and metapopulation IPMs estimated different parameterizations of movement dynamics (i.e., movement was ignored, estimated as a time- and age-invariant rate, or estimated as time-invariant and age-varying), and incorporation of alternate data (i.e., with or without fitting tag-recapture data; Figure 1, Table 1). The robustness of each IPM to the various true spatial dynamics of the three OMs was demonstrated by calculating estimation bias and associated metrics for important conservation parameters used for providing management advice (e.g., spawning biomass, fishing mortality, and recruitment). The modeling framework is described below with a focus on spatial dynamics and population structure differences among OM and IPM parametrizations. A detailed description of the assumed life history dynamics, input population parameters, and rationale for the OM are provided in the Supplementary Material. All models were developed in AD Model Builder (Fournier et al., 2012) with visualization and performance metrics calculated in R (R core team, 2018). Each version of the OM and IPM can be downloaded from the GitHub repository (https://github.com/KateBoz/Spatial_IPM).

2.2 Operating Model

The OM structures were based on inputs and results from the most recent sablefish stock assessment (Hanselman et al., 2018), recent analysis of tag-recapture data (Hanselman et al., 2015), and feedback from the development of a spatial IPM (Fenske, personal communication). Input parameters either came directly from the most recent stock assessment or were structured using hypotheses and ongoing research regarding the spatial population dynamics of sablefish. We first describe the common dynamics across model types then separately describe the population structure and parameterization of the three OMs (*Uniform*, *Spatial Heterogeneity* (*SH*), and *Metapopulation*).

2.2.1 Operating Model Dynamics

The sablefish OM consisted of three areas across which sablefish could be homogeneous or heterogeneously distributed, depending on assumed spatial structure and connectivity dynamics (Fig. 1). Population abundance by year (y) and age (a) was projected forward using population dynamics equations where the sequential order of events in a given yearly time step involved: (1) spawning; (2) recruitment; (3) release of tagged fish, if tagging takes place; (4) instantaneous movement of tagged and untagged fish between areas; and (5) continuous natural mortality and harvest throughout the year, including tag recaptures. Abundance (N) of fish in a

given area (p; commensurate with a subpopulation in the metapopulation models since fish immediately assume the biological parameters of a new subpopulation upon moving) was projected forward for 30 years starting from an assumed initial abundance-at-age. Initial abundance of age-1 individuals (i.e., recruits) was either equal to the area-specific average recruitment (R_{ave}) parameter for the *Metapopulation OM* or the total system-wide average recruitment multiplied by the area-specific recruit apportionment parameter (ξ) for the *Uniform* and *SH* OMs (see Table 2 for parameter values). Initial abundance of ages 2 through the plus group (age 16+) was calculated as an exponential decay from initial age-1 abundance based on the natural mortality rate ($M = 0.1$ for all ages and areas).

After the first year, abundance-at-age was calculated at the beginning of the year (y) before movement occurred (N_{BEF}) based on the abundance after movement (N_{AFT}) in the previous year and age and discounted for natural and fishing (F) mortality:

$$N_{p,y,a,BEF} = N_{p,y-1,a-1,AFT} e^{\left[-\left(F_{p,y-1,a-1}+M\right)\right]} .$$

Eqn. 1

Fully selected fishing mortality assumed a dome shape across the time series, where it increased linearly from a specified minimum (F_{min}) in the first year to a specified maximum halfway through the time series (F_{max}), then decreased linearly again through the end of the time series (see Table 2 for input values). Annual lognormal deviations by area, defined by an input variance term (σ_F), were applied to mimic random noise in the fishing process. Values for F_{min}, F_{max}, and σ_F were allowed to vary by area in the spatially explicit OMs (Tables 2 and 3). Fishery selectivity (v_f; susceptibility to the fishing gear) was modeled with a two-parameter logistic function (Fig. B2) and was area- and time-invariant. The total fishing mortality on a given age was the combination of selectivity-at-age and fully selected fishing mortality by year.

When connectivity occurred (i.e., all models except the *Uniform* OM) the box-transfer method was utilized, which assumed movement was a Markov process. The movement parameter, $T_{y,a}^{j\to p}$, represented the fraction of age a fish from area j in year y that moved to area p. Abundance after movement was given by:

$$N_{p,y,a,AFT} = \sum_{j=1}^{p} \left[T_{y,a}^{j\to p} N_{j,y,a,BEF}\right] .$$

Eqn. 2

The *Uniform* OM assumed no movement among areas whereas the *SH* and *Metapopulation* OMs included both age- and time-varying connectivity patterns. Age-specific movement rates were

derived from length-based estimates of sablefish movement within Alaskan waters determined from analysis of tag and recovery data (Hanselman et al., 2015). Simulated movement rates were binned into three age blocks (ages 1-4, 5-9, and 10-16), which differed by area (Table 4). Annual deviations that varied by area were applied to the age-specific movement rates to generate time-varying movement. Annual movement deviations increased in 10-year time blocks to mimic increasing variability in movement over time (σ_T = 0.02, 0.04, 0.06 for each time block, respectively; see Fig. B4 for an example of the connectivity patterns simulated).

Spawning stock biomass (*SSB*), a measure of potential population productivity, at the beginning of the year was the product of abundance, input area-specific maturity (*m*; except for the *Uniform* OM, which used the Area 2 values for all areas; Fig. B3), and input area-invariant weight (*w*; Fig. A2):

$$SSB_{p,y} = \sum_{a=1}^{A} N_{j,y,a,BEF} w_a m_{p,a}$$

Eqn. 3

New births or recruitment were based on an area-specific input average recruitment term, $R_{ave,}$ multiplied by an area-specific recruitment apportionment with bias corrected lognormally distributed area-specific annual random deviations (ε_R) controlled by the area-invariant recruitment variance term (σ_R = 0.9):

$$N_{p,y,a=1,BEF} = \xi_p R_{ave,p} e^{\left(\varepsilon_{R_{p,y}} - 0.5\sigma_R^2\right)}; \varepsilon_{R_{p,y}} \sim N\left(0, \sigma_R^2\right) .$$

Eqn. 4

In the *SH* and *Uniform* OMs, a single stock-recruit relationship and associated R_{ave} value was utilized where recruitment deviations were applied at the global level (i.e., not area-specific) and were identical among these OMs. Area-specific recruitment was determined by applying the recruitment apportionment term with equivalent apportionment assumed for each area in the *Uniform* OM and spatially varying apportionment for the *SH* OM (Table 2). For the *Metapopulation* OM, each individual subpopulation was assumed to have its own stock-recruit relationship where the average recruitment parameters were subpopulation-specific and no apportionment occurred (i.e., the recruit apportionment term was set to 1.0 within each subpopulation). However, the population-specific average recruitment terms of the *Metapopulation* OM were scaled to the area-specific recruit apportionment terms of the *SH* OM

307 to maintain relative consistency in recruit dynamics among these models (see Table 2 for input
308 parameters). For the *Metapopulation* OM, the recruit deviations differed by area, but recruitment
309 variance terms were area-invariant (identical to the other OMs).

310 ***2.2.2 Operating Model Population Structures***

311 Different OMs were constructed to represent a range of complexity in population
312 structure, which emulated those most commonly observed for marine fish populations (see
313 Goethel et al., 2011 and Cianelli et al., 2013). Specific dynamics for each of the three OM
314 configurations are provided below.

315 *Uniform OM*

316 The *Uniform* OM emulated the dynamics of a homogeneous (i.e., panmictic) population
317 distributed evenly across the three areas. By simulating the dynamics in three areas even though
318 they were identical, it allowed pseudo-data to be provided by area (although the data were also
319 identical) and enabled the application of both spatially aggregated (i.e., *Panmictic*) and spatially
320 explicit IPMs. The *Uniform* OM assumed all parameters were identical across areas (the
321 parameter values for each area matched those of Area 2 applied in the spatial models; Table 2),
322 while movement did not occur among areas. The *Uniform* OM also assumed a single
323 reproductive unit where annual population-level recruitment was apportioned equally among
324 areas. Fishing mortality rates were assumed identical across areas (Table 2). Tagging data were
325 not simulated, because movement did not occur and the role of tag-recapture data in this study
326 was to help estimate movement rates.

327 *Spatial Heterogeneity (SH) OM*

328 The *Spatial Heterogeneity* (*SH*) OM was configured to simulate a single population with
329 spatial heterogeneity across each area, which was created through spatial variation in
330 demographics, fishery dynamics, and connectivity. Similar to the *Uniform* OM, the population-
331 level recruitment was apportioned to each area, but varied among the areas (i.e., ξ= 0.44, 0.30,
332 and 0.26 for Areas 1, 2, and 3, respectively). Additional heterogeneity was created by allowing
333 maturity ogives to vary among areas (Fig. A3). The annual fishing mortality rate also varied
334 spatially with differing specifications for F_{max}, F_{min} and σ_F (Tables 2 and 3). Connectivity
335 dynamics were simulated as time and age-varying following the parameterization described
336 below (Table 4). Yearly tagging data were simulated, but only fit in IPMs scenarios that included
337 tagging data.

Metapopulation OM

The *Metapopulation* OM was designed to simulate three subpopulations each occupying one of the three areas, which were connected through movement. A fish moving among areas was assumed to adopt the biological characteristics of the new subpopulation immediately upon entering a new area. Each subpopulation had its own stock-recruitment relationship with spatially varying R_{ave} values and recruitment deviations. Subpopulation specific R_{ave} values were specified such that the age-1 initial abundance in each subpopulation matched the respective area-specific values from the *SH* OM configuration. This allowed continuity in recruitment levels among the different OMs. Apart from the recruitment dynamics and population structure, all other parameters were assumed to be identical to those from the *SH* OM. Once again, yearly tagging data were simulated, but were only fit in IPM scenarios that included the tagging data.

2.2.3 Data Generation

Each OM generated simulated pseudo-data typical to IPMs used in fisheries applications. Simulated area-specific data sources included landings (i.e., total biomass of landed catch), age composition of the landed catch, a fishery independent survey of biomass, age composition of the survey biomass, and, for certain scenarios, tag-recapture data. The fishery was assumed to operate continuously for the entire yearly time step and area-specific catch was calculated using Baranov's catch equation (Baranov, 1918) based on the area-specific fishing morality and abundance. The fishery-independent survey (s) was assumed to occur mid-year ($t_s = 0.5$) where area-specific survey catch, discounted for mortality up to the time of the survey, was calculated using the same approach as the fishery catch. The survey assumed a time- and area-invariant 2-parameter logistic survey selectivity function (v_s; Fig. A2) and a time- and area-invariant survey catchability scalar. The *SH* and *Metapopulation* OMs also simulated tag release and recapture data using a multiyear Brownie tag-recovery model (Brownie et al., 1993). In each year of the simulation, a new tag cohort was released into the population, where a cohort (l) was defined by the combination of year, age, and area of release. The total number of tag releases in each year was based on a specified tag proportion parameter ($\rho = 0.0005$), which proportionally scaled tag releases by the total survey abundance. Annual tag releases were then distributed across areas based on relative survey abundance in each area and across ages based on survey selectivity. Tag abundance (n) by cohort was calculated similar to the main population (i.e., following Equations 1-2), but with recruitment replaced by tag release events. Cohort specific recaptures (r) were

calculated using Baranov's catch equation assuming 100% tag reporting (β = 1.0; 100% tag reporting was assumed for model simplification):

$$r^{l}_{p,y,a} = n^{l}_{p,y,a,AFT}\frac{F_{p,y,a}\left(1 - e^{\left[-\left(F_{p,y,a}+M\right)\right]}\right)}{F_{p,y,a}+M}.$$

Eqn. 5

Measurement error for each data source was simulated using stochastic processes based on an assumed underlying probability distribution (Table B1), which resulted in the final 'observed' pseudo-data that were eventually fit within each IPM. For each simulation scenario, 150 stochastic simulations were conducted where each iteration generated a unique 30-year time series of pseudo-data from the OM. All scenarios used the same vector of randomly generated seeds for the 150 simulations. The assumed probability distribution (lognormal or multinomial) and associated error level (input variance or effective sample size, *ESS*) are provided in Table 3. The error levels and number of runs were chosen to adequately encapsulate stochasticity and represent average variation often assumed for marine data collection programs. A multinomial probability distribution was utilized for the tagging data, but the *ESS* was set at 200, which was lower than the actual number of tags released per cohort. The lower *ESS* increased uncertainty (i.e., allowed for implicit overdispersion) in the tagging data. Otherwise, the tagging data would have been overly informative compared to real-world data collection. **2.3 Integrated Population Models**

Four versions of an IPM were developed to evaluate the impact of incorrect assumptions regarding population structure and movement dynamics (Table 1). The suite of IPMs tested included *Panmictic*, *Fleets-as-areas (FAA)*, *Spatial Heterogeneity (SH)*, and *Metapopulation* models. The underlying population dynamics equations and specifications for the IPMs matched the corresponding OMs, except for the spatially aggregated models (i.e., the *Panmictic* and *FAA* IPMs; specific differences are outlined in the following sections on IPM spatial structure).

2.3.1 IPM Spatial Structure

Panmictic IPM

The *Panmictic* IPM assumed a single homogenous population across the entire model domain (i.e., a one area model with no movement), which likely represents the most common approach to fisheries stock assessment (i.e., assuming a closed unit population; Punt, 2019a, b) and the current method applied in IPMs for sablefish. In the *Panmictic* IPM, parameters were

399 estimated as a unit population (i.e., area-specific values were not estimated). Area-specific data
400 sources and inputs from the OMs were additively combined (e.g., fishery yield and survey
401 biomass) or aggregated as abundance-weighted averages (i.e., fishery and survey age
402 compositions, weight-at-age, and maturity-at-age) and fit in the IPM at the aggregated scale.
403 Tagging data were not fit in the *Panmictic* IPM and movement rates were not estimated, because
404 only one spatial area was assumed to exist.
405 *Fleets-as-Areas (FAA) IPM*
406 The *FAA* IPM assumed a single population with no explicit spatial structure, but spatially
407 varying fishery parameters were estimated (by fitting spatially disaggregated data from these
408 fleets) to implicitly account for spatial dynamics. *FAA* IPMs are often implemented when little or
409 no information on spatial structure exists, but spatially disaggregated fishery data are available.
410 Modeling the spatial variability in the fishing fleets serves as a proxy for the actual spatial
411 structure without needing to make assumptions about the underlying population structure and
412 avoiding the need to explicitly model connectivity. For our study, the *FAA* IPM had fixed
413 recruitment apportionment that was equal among areas (i.e., set at 0.33 per area), identical
414 biological parameters across areas (akin to the *Panmictic* IPM), and assumed no movement
415 occurred among areas. Fishery selectivity and fishing mortality were estimated by area. One
416 survey selectivity and one catchability were estimated. Tagging data were not fit in the *FAA*
417 IPM. The *FAA* IPM was utilized to determine if underlying spatial variation in the population
418 could be effectively captured through area-specific fishery parameter estimates without needing
419 to implement a spatially explicit IPM.
420 *Spatial Heterogeneity (SH) IPM*
421 The *Spatial Heterogeneity* (*SH*) IPM mirrored the parameterization of the *SH* OM
422 allowing spatial heterogeneity within a single population unit by explicitly accounting for spatial
423 variation by area. A single stock recruitment function (i.e., one R_{ave} parameter) was estimated
424 with unequal recruitment apportionment fixed at the values from the *SH* OM. Fishery selectivity,
425 fishing mortality, and movement rates were estimated as area-specific. One survey selectivity
426 and one catchability were estimated. Depending on the model scenario, movement rates were
427 estimated to be either constant values (i.e., age- and time- invariant) or, if tagging data were fit,
428 age-varying. Tagging data were fit in the *SH* IPM for several scenarios, but not all (see Table 1
429 for all scenarios and model parameterizations). We considered the performance of the *SH* IPM to

represent a best-case representation example for most scenarios, because the apportionment parameters were either fixed at the true value (when applied to outputs from the *SH* OM) or directly matched the spatial distribution of R_{ave} (when applied to the outputs from the *Metapopulation* OM). We also fit the *SH* IPM with no movement estimated and fixed at 100% residency to the *SH* and *Metapopulation* OMs. This approach is used when regional population structure is identified, but there is limited knowledge of connectivity among population units.

Metapopulation IPM

The *Metapopulation* IPM matched the structural assumptions of the *Metapopulation* OM assuming three subpopulations connected through post-settlement movement. The *Metapopulation* IPM estimated area-specific values for average recruitment, recruitment deviations, fishing mortality, fishery and survey selectivity, and survey catchability. In addition, movement was estimated between areas. Depending on the model scenario, movement rates were assumed to be either constant (i.e., age- and time-invariant) or, if tagging data were included, age-varying. Tagging data were fit in the *Metapopulation* IPM for several scenarios, but not all (see Table 1). The *Metapopulation* IPM is the most spatially complex IPM tested, and it emulates the population structure most widely hypothesized for marine species (Smedbol and Stephenson, 2001; Goethel et al., 2011).

We also explored a closed population parametrization of the *Metapopulation* IPM where movement was not allowed among populations (and was not estimated) and no tagging data were fit in the IPM. This approach is often suggested as the first step towards developing fully spatial IPMs (Cadrin, 2020). Each area assumed a unit population with a unique stock recruit relationship. The three independent, closed populations were modeled simultaneously with spatially varying parameters estimated for each population. All parameters were estimated as area-specific including average recruitment, recruitment deviations, fishing mortality, selectivity, and catchability.

2.3.2 Integrated Population Model Estimation

Estimated parameters for each IPM included survey catchability, annual fishing mortality rates, average recruitment (R_{ave}), annual recruitment deviations, and logistic parameters for survey and fishery selectivity. For some IPMs, these quantities were also estimated as area specific. In addition, connectivity among areas was directly estimated for the *SH* and *Metapopulation* IPMs depending on the scenario being tested. A multinomial logit

transformation was utilized for movement parameters to naturally bound parameters between zero and one and to ensure that the summation of emigration and residency equaled unity for a given population. Only the off-diagonal elements (i.e., emigration rate from an area) of the movement matrix were estimated to ensure identifiability of the model, while the diagonal elements (i.e., residency) were calculated as one minus the sum of emigration from a population. Movement was treated as time-invariant resulting in a total of 6 emigration parameters (two per area) to be estimated. Depending on the scenario, age-varying movement could also be estimated, which resulted in a total of 96 estimated movement parameters. Natural mortality and recruit apportionment (where applicable), as well as growth and maturity were fixed at the true values from the OM. These values (e.g., natural mortality) were fixed at the true values to reduce the number of estimated parameters, and because they are commonly fixed in fishery IPMs.

Parameters in the IPMs were treated as fixed effects and estimated with a maximum likelihood (MLE) framework, which integrates numerous data sources, through an objective function, and allows each data source to assume a specified underlying error structure (Maunder and Punt, 2013). The IPMs minimized differences between model predicted observations and the pseudo-data generated from the OM for each data source. The total likelihood was determined by summing the negative logarithm of each likelihood component, which was then minimized to derive best fit parameter estimates. Data used to calculate the individual likelihood components and associated assumed distributions were fishery landings (lognormal); survey biomass (lognormal); fishery age compositions (multinomial); and survey biomass age compositions (multinomial). Tag recapture proportions (multinomial) were also included for scenarios that estimated movement using tagging data.

$$-ln(L_{total}) = -\ln(L_{F\_yield}) - \ln(L_{S\_bio}) - \ln(L_{F\_comp}) - \ln(L_{S\_comp}) - \ln(L_{Tag\_rec}). \quad \textbf{Eqn. 6}$$

MLE variance terms for each likelihood component were taken directly from the operating model except for the recruitment variance where the IPM assumed a larger variance than was used for data generation. Similarly, the effective sample size for multinomial distributions was reduced by 100 for each data source to avoid overfitting age composition and tagging data (see Table B1 for input error terms). Penalty functions were used to stabilize estimates and prevent unfeasible parameter values (e.g., zero values of average recruitment; extremely high large movement, fishing mortality, or recruitment deviations).

2.4 Evaluation of Model Performance

Model performance was evaluated by calculating the bias and precision of estimated parameters from converged model runs, with primary focus on spawning biomass, recruitment, and fishing mortality rates. Convergence criteria included the ability to calculate a positive-definite Hessian matrix and having a maximum objective function gradient less than 0.001. The convergence rate across the 150 simulated iterations within a scenario provided a measure of model stability. Relative error level of a specific parameter for a given year (y), area (a) and scenario (k) was evaluated based on the relative percent difference (*RPD*) between the estimated parameter ($\bar{x}$) for a given model iteration (z) and the true value used in the OM (x), such that:

$$RPD_{k,a,y,z} = \left(\frac{\bar{x}_{k,a,yz} - x_{k,a,yz}}{x_{k,a,y,z}}\right) \cdot 100$$

The $medianRPD$ was then calculated per year (y) and area(a) across iterations for a given scenario (k).

$$medianRPD_{k,a,y} = median(RDP_{k,a,y,1} ... RDP_{k,a,y,150})$$

Eqn. 7

An aggregated relative error metric, the scaled cumulative absolute percent error (which we termed *sCAPE*), was developed to evaluate the overall bias and precision of a parameter for each scenario and to compare performance of the IPMs when provided data from different underlying spatial population structures. The *sCAPE* metric first calculates the cumulative sum of the *medianRPD* absolute values across a time series for a given area. The cumulative absolute percent error, *CAPE*, is then scaled to the maximum *CAPE* value across all scenarios for a given quantity to produce an area specific *sCAPE*.

$$CAPE_{k,a} = \sum_{y=1}^{30} (|\, medianRPD_{k,a,y} \,|)$$

Eqn. 8

$$sCAPE_a = CAPE_{k,a} \cdot 1/\, max(CAPE_k)$$

Eqn. 9

The *sCAPE* metric provides a measure of IPM performance, with values closest to zero indicating greater accuracy in an estimated parameter relative to all other scenarios. The *sCAPE* metric was used to compare across IPM types whereas a *cumulative sCAPE* was used to compare

performance within IPM types given different underlying spatial population structures and model parameterizations. The cumulative *sCAPE* summed all area specific values into a single error metric with values closest to zero indicating improved model performance.

$$\text{cumulative sCAPE} = \sum_{a=1}^{3} sCAPE_a$$

Eqn. 10

Figures showing the distribution of parameter estimates and *RPD* values for all model iterations across the time series were used to evaluate the magnitude and direction of parameter bias within a given scenario. In addition, the distribution of RPD values associated with terminal year *F* and *SSB* for all scenarios was examined, because these values represent important quantities used to inform fisheries management. Where possible, parameter estimates were provided by area and for the entire system (except for fishing mortality, because it is not straightforward to aggregate area-specific instantaneous rates to a system level rate when different estimates of selectivity exist for each area). Additionally, the *sCAPE* values were used to evaluate performance of spatially explicit IPM across the entire complement of simulated spatially explicit OM population structures. Given that the *Panmictic* and *FAA* IPMs could not estimate area-specific parameter values for all parameters, these IPMs were compared using only the system level *sCAPE* metric for each parameter. The best performing IPMs were those with configurations that had the smallest *sCAPE v*alues across all OMs. The *cumulative sCAPE* solution provided guidance on which population structure parametrization of a given IPM was most robust to uncertainty in true underlying population structure.

2.5 Scenarios and Sensitivity Runs

We simulated three OM parametrizations (i.e., *Uniform*, *SH*, and *Metapopulation*) and used four spatial structure assumptions in the IPMs (*Panmictic*, *FAA*, *SH*, and *Metapopulation*). We also used three configurations for the spatially explicit IPMs to account for movement dynamics (i.e., estimating time- and age-invariant movement, estimating age-varying movement, or a closed population [no movement]) and two data configurations for the spatially explicit OMs to account for data availability (i.e., assuming no tagging data were available or directly fitting tagging data in the objective function). A full factorial design was implemented where each parametrization of the spatial structure in the IPM was applied to the data generated from each of the potential spatial structures in the OM (see Table 1 for a complete list and associated scenario

551 names). Each scenario is referred to by the following convention: *OM:IPM:Movement:Tags*.
552 The combinations allowed for a relatively complete comparison of how the most widely applied
553 spatial population assumptions in IPMs performed with no *a priori* knowledge of the underlying
554 true spatial population structure. Exploration of bias in these IPMs provides a demonstration of
555 how well they might be expected to perform in real-world applications when developing
556 management advice, whereas the *cumulative sCAPE* solution provides initial evidence for which
557 IPM parametrizations could be the most robust given population structure uncertainty.

558 **3 Results**

559 Most IPMs achieved near 100% convergence (Table 1). The lowest convergence rate was
560 91% for the *SH* IPM applied to the *SH* OM with age-based movement estimated and tagging data
561 included. These high convergence rates generally indicate that the models were relatively stable
562 with limited overparameterization and no extreme parameter correlation.

563 Overall IPM performance differed based on the spatial structure of the OM and
564 parameterization. Generally, IPM models that estimated movement and included tagging data
565 were robust to mismatch in assumed spatial structure (Table 3; Figure 2). Generally, when the
566 IPM structure matched that of the OM, the matching IPM tended to provide the lowest *sCAPE*
567 values for all parameters compared to mismatched IPMs. Similarly, terminal year estimates of
568 fishing mortality and *SSB* were generally more accurate and precise when the IPM and OM
569 structures matched (Figure 3), as would be expected. At the system level, most of the
570 combinations of IPM and OM provided unbiased estimates of the terminal year *SSB* even when
571 population structure assumptions were mismatched; however, the individual estimates by area
572 were biased in some scenarios (Figures 2 and 3). The terminal year system level *F* and *SSB* were
573 unbiased for the *Panmictic* IPM for all OMs except the *Metapopulation* IPM and for the *FAA*
574 IPM with the *Uniform* OM (Figure 3). In general, the estimation of *SSB* was more accurate than
575 the estimation of *F*. The largest bias in the spatially explicit IPMs occurred when a constant
576 movement rate was estimated with or without tagging data or when no movement was estimated
577 (Figure 2 and 3). The best performing spatial models were those that allowed for the estimation
578 of age-based movement and incorporated tagging data (Figures 2 and 3).

579 **3.1 Panmictic IPM**

580 The *Panmictic* IPM was relatively robust to the underlying population structure for
581 estimating system level parameters. The *Panmictic* IPM had low cumulative *sCAPE* values

582 across the OM population structures with values ranging between 0.002 and 0.119 (Table 3). The
583 *Panmictic* IPM performed best for the *Uniform* OM and demonstrated only slight bias in system
584 level estimates of terminal year *SSB* yet had increased bias in the estimates of terminal year
585 fishing mortality rate when mismatched to the *SH* and *Metapopulation* OMs. When mismatched
586 to the underlying population structure, the *Panmictic* IPM demonstrated strong directional bias at
587 the beginning of the time series, but with decreasing, yet variable, bias towards the end of the
588 time series (Figure 4).

589 **3.2 Fleets-as-Areas IPM**

590 Overall, the *FAA* IPM performed well when the underlying population was uniform, but
591 performed poorly when underlying spatial dynamics were present. The *FAA* IPM had low
592 cumulative *sCAPE* values for the *Uniform* OM with values ranging between 0.003 and 0.036
593 (Table 3) and was able to estimate terminal year *SSB* and area specific fishing mortality rates
594 with no bias for this OM (Figures 2, 3, and 4). In contrast, the *FAA* IPM had higher cumulative
595 *sCAPE* values for the *SH* and *Metapopulation* OMs with values ranging from 0.023 to 1.124
596 (Table 3) with highly biased estimates for the terminal year *SSB* and area specific fishing
597 mortality rates (Figures 2 and 3). Although the *FAA* IPM demonstrated limited bias in *SSB* at the
598 beginning of the time series when the OM assumed a spatially explicit population structure, bias
599 increased dramatically and unidirectionally as time progressed (Figure 4).

600 **3.3 Spatial Heterogeneity IPM**

601 The *SH* IPM was generally robust to underlying population structure but performed best
602 when allowed to estimate age-based movement or when the underlying population structure was
603 uniform (Table 3). On both a system level and for each population unit, estimating age-based
604 movement with tagging data gave the least biased results for both the *SH* and *Metapopulation*
605 OMs (Figures 5 and 6). When the *SH* IPM was matched to the *SH* OM but estimated a constant
606 movement rate without including tagging data, it led to a lower cumulative *sCAPE* value for *SSB*
607 and recruitment estimation compared to the same configuration with tagging data included
608 (Table 3; Figure 2). Conversely, not including tagging data led to higher cumulative *sCAPE*
609 values for *F* and movement estimation. Not estimating movement had higher sCAPE values and
610 biased estimates of terminal year SSB and fishing mortality rate (Figures 2 and 3). For the
611 *Metapopulation* OM, estimation of constant movement with and without tagging data did not
612 cause much difference in the cumulative *sCAPE* values, yet not estimating movement at all

resulted in much higher cumulative *sCAPE* values (Table 3). Although the *SH* IPM mismatched to the *Metapopulation* OM performed well for most parameters when age-based movement was estimated, it was unable to accurately estimate recruitment due to the fixed recruit apportionment parameters and a single set of recruit deviations (i.e., compared to area-specific stock-recruitment curves and area-specific recruitment deviations assumed in the *Metapopulation* OM). Overall, incorrectly specifying movement was more detrimental than incorrectly specifying the underlying population structure.

3.4 Metapopulation IPM

The *Metapopulation* IPM performed very similarly to the *SH* IPM with the best performance occurring when age-based movement was estimated or when the underlying population structure was uniform, as indicated by the lowest cumulative *sCAPE* values for these scenarios (Table 3). When the *SH* OM was used, the *Metapopulation* IPM that estimated a constant movement rate while including tagging data had lower *sCAPE* values for *SSB, F,* and movement compared to the same configuration without tagging data (Figure 2). On both a system level and for each population unit, estimating age-based movement with tagging data gave the least biased results for both the *Metapopulation* and *SH* OMs (Figures 7 and 8). Bias increased for both the system-level estimates and for each area when a constant movement rate was estimated or movement was not estimated when both the *Metapopulation* and *SH* OMs were applied. The *Metapopulation* IPM generally had relatively high *sCAPE* values in estimating area-specific recruitment, likely due to the added parameters that needed to be estimated for area-specific stock-recruitment curves and associated deviations (Figure 2).

4 Discussion

By developing a spatially explicit simulation-estimation framework and exploring a variety of population structure and movement assumptions, we were able to demonstrate the general robustness of spatially explicit IPMs using Alaskan sablefish as a case study. Regardless of the underlying population structure (including an essentially homogeneous stock with no movement, i.e., the *Uniform* OM), each of the spatially explicit IPMs were able to accurately estimate area-specific parameter values and increase precision when flexible parameterizations of movement were utilized (i.e., age-based) and auxiliary tagging data were applied (See Table B3 in the Supplementary Material). The *Panmictic* IPM was generally robust to underlying spatial structure when estimating system level parameters but would provide no support for

developing area-specific management advice. Conversely, spatially implicit IPMs (i.e., *FAA* IPM) provided area-specific fishing mortality, but estimates were generally biased when confronted with underlying spatial population structure. Therefore, the results of this study indicate that when underlying population structure is likely to be present and spatial management is needed (i.e., to protect subpopulation or spawning components or to control spatially variable harvesting or fleet structure), then spatially explicit IPMs should be utilized that incorporate enough estimation flexibility to emulate important drivers of spatial dynamics.

Our results provide further support for the general findings that suggest that spatial IPMs are likely to be more robust than spatially aggregated or panmictic IPMs even when limited understanding of underlying spatial dynamics exist (e.g., Ying et al., 2011; Goethel et al., 2015a, 2021; Punt, 2019a,b). Although assuming a unit population provided unbiased estimates of system level parameters, the potential for localized depletion when subpopulation structure is ignored has been widely acknowledged (Fu and Fanning, 2004; Ying et al., 2011). On the other hand, Punt et al. (2018) demonstrated that assuming highly complex spatial dynamics was less detrimental than implementing simplified models. Our results support this conclusion and demonstrate that allowing for spatial population structure is likely to be less detrimental than ignoring it completely. Furthermore, our study clearly illustrates that allowing for flexibility in the parametrization of movement is more important than correctly specifying spatial population structure. When the *SH* and *Metapopulation* IPMs estimated age-varying movement, the outputs were essentially unbiased despite the potential for incorrect assumptions regarding population structure. Ignoring age-based movement in the spatially explicit IPMs led to biased area-specific parameter estimates. These results support previous research (Ying et al., 2011; Goethel et al., 2015b; Lee et al., 2017; Cadrin et al., 2019; Goethel et al., 2021), which suggests that simplified movement dynamics can be as detrimental to spatial IPM performance as ignoring movement altogether. Estimating the full complexity of movement is intractable and movement dynamics are often as uncertain as population structure. Goethel et al. (2021) suggest using flexible movement parameterizations that balance parsimony and complexity, while focusing on estimating along the axis that is most likely to drive spatial dynamics for the given species. Combined with the use of random effects to help estimate time-variation in recruitment and movement parameters (Thorson et al., 2015), flexible movement parametrizations implemented

within spatially explicit IPMs are likely to allow these models to provide robust outputs that can adequately support spatial management measures.

When little is known regarding spatial dynamics in marine resources, the first step should always be to perform a holistic stock identification study (e.g., Cadrin et al., 2014; Cadrin, 2020) to identify the spatial scale of important population components that require monitoring and independent management. The management and stock assessment boundaries should then be adjusted to match these units (Kerr et al., 2017; Cadrin, 2020). Although implementing closed population IPMs on these units is often touted as the next step towards developing full spatial IPMs and is sometimes adequate when limited movement exists (e.g., Cadrin et al., 2019; Goethel et al., 2015a,b), our results suggest that there is limited cost to implementing a full spatial model even if population structure and movement are not fully understood. Conversely, ignoring movement in the closed population models led to high levels of bias, which supports the findings of Ying et al., (2011) where closed population models were shown to lead to overexploitation of subpopulations within a metapopulation.

Accurate estimation of movement parameters or mixing among populations often requires additional data sources, such as tagging or genetic data (Vincent et al., 2017; Goethel et al., 2019, 2021). For our model, tagging data improved the estimation of movement with increased precision and accuracy in parameter estimates, even when population structure was mis-specified. When tagging data were not available, spatial IPMs are still able to estimate movement (e.g., Hulson et al., 2011, 2013; McGilliard et al., 2015). However, imprecision in parameter estimates often increases drastically, confounding with recruitment parameters may occur, and estimation of more complex movement patterns becomes difficult (Goethel et al., 2019). Incorporating traditional tagging data may also be problematic if information on tag reporting rate, tag mixing, or the age or length structure of the released and recaptured tags is not well known (Goethel et al., 2019). We assumed 100% tag reporting and no tag loss, which would not be the case for empirical tagging data and likely produced optimistic estimates of parameter bias. In data limited situations when no additional information is available to inform movement rates or when tagging data are likely to be unreliable, closed population IPMs applied at the scale of important subpopulation components should be considered (Goethel et al., 2015b; Cadrin et al., 2019). Several approaches are available to deal with the assumptions of traditional tagging data (Goethel et al., 2019). Advances in electronic tagging, genetic methods, and remote sensing

technology (e.g., drones and satellite imaging) have led to a proliferation of data that has identified migration corridors, movement patterns, and mixing rates among population units for a wide variety of species (Bravington et al., 2016; Lowerre-Barbieri et al., 2019). New and evolving data types combined with the power of integrated analyses allow spatially disparate data sets to be combined into a single spatially explicit IPM to estimate shared or spatially distinct parameters, suggests that spatial IPMs should be more widely applied (Berger et al., 2017; Zipkin and Saunders, 2018; Goethel et al., 2021).

Alternate spatially explicit data sources can also aid in the estimation of area or population specific recruitment in spatial IPMs. Movement and recruitment estimates are often highly correlated in spatial IPMs (Cadrin et al., 2019), and our results demonstrated that the addition of tagging data reduced bias in recruitment estimates for the spatially explicit IPMs. However, results from the *SH* IPM are likely to be optimistic, given that the recruit apportionment parameters were fixed. Because the fixed values matched the relative split of average recruitment across subpopulations in the *Metapopulation* OM, it is likely that the *SH* IPM performed excessively well when confronted with the *Metapopulation* OM. Although exploratory runs attempted to estimate recruit apportionment, most runs failed to converge and were excluded from the full analysis. Recruit apportionment models are widely applied (e.g., generalized assessment framework Stock Synthesis 3 uses the *SH* IPM approach described in this paper; Methot and Wetzel, 2013) and a variety of methods exist for implementation (e.g., fixing apportionment parameters, estimating time-invariant values, or estimating time-varying values; Punt, 2019a). Although similar spatial simulation studies have shown limited bias when apportionment has been estimated (e.g., Punt et al., 2015, 2019a; Denson et al., 2017), these models rarely combine the estimation of complex movement and recruit apportionment. Other studies have taken a similar approach to our study and fixed the recruit apportionment parameter when complex spatial dynamics were modeled (Little et al., 2017), acknowledging that the actual bias is likely to be much higher in real-world applications when recruit apportionment is fixed. Future studies should consider further exploration of the performance of the *SH* IPM framework, especially when combined with complex movement estimation.

Although all IPM configurations tested demonstrated relatively poor performance when confronted with certain OM configurations, the *FAA* IPM performed consistently poorly when confronted with spatially explicit OMs. Bias was limited at the beginning of the timeseries but

accumulated over time. In particular, when the *FAA* IPM was supplied data from the *SH* OM, the mismatch in parameterization of recruitment apportionment and movement caused bias in annual recruitment (which was over-estimated) and fishing mortality (which was underestimated). Given that the sole purpose of implementing a *FAA* IPM is to implicitly account for spatial structure by modeling unique fishing fleets in each spatial area, our results suggest that the *FAA* IPM should not be utilized. Again, this supports recent suggestions that FAA approaches are generally not advisable (Hurtado-Ferro et al., 2014; Punt et al., 2016; Punt, 2019b). However, *FAA* IPMs can sometimes outperform spatially naïve IPMs (Lee et al., 2017), and thus may be useful when no additional spatial data are available, little is known about the spatial dynamics present, and complex fleet structure exists. Depending on the management need and complexity of available data, a *FAA* model may provide perform adequately and produce outputs at the desired spatial scale. Before implementing a *FAA* IPM, though, thorough vetting and simulation testing should be undertaken to ensure that a spatially implicit model is indeed likely to outperform a *Panmictic, SH,* or *Metapopulation* IPM.

The operating models developed for this study represent some of the most spatially complex OMs that have been used to simulation test IPMs, because they were meant to emulate the complex real-world spatial dynamics of sablefish. Despite the multiple spatial complexities included (e.g., complex population structure, recruitment dynamics, and age- and time-varying movement), the simulation models were still relatively simplified compared to what would be expected in a real-world application. In particular, the level of misspecification for some processes in the applied IPMs is much lower than would be expected given that many parameters (e.g., *M*, weight, and maturity) were fixed at their true values. Additionally, it was assumed that the system was completely closed to immigration or emigration and that the area boundaries were accurately represented (i.e., the boundaries correspond exactly with subpopulation components and the extent of the associated fishery). Therefore, these results are expected to be extremely optimistic. If increased misspecification were present or individuals were migrating outside of the system boundary, increased bias would be expected (Berger et al., 2020).

Many aspects of spatial IPM performance remain to be explored before these modeling approaches are more generally adopted as the basis of fisheries management advice worldwide (Berger et al., 2017; Punt, 2019b). We have demonstrated that the assumption of spatial heterogeneity and metapopulation spatial structure appears to be relatively robust to incorrect

767 specification in a spatial IPM. However, our analysis did not explore whether these assumptions
768 are robust to natal homing, another widely observed spatial populations structure with unique
769 spatial dynamics (e.g., strong natal fidelity, directed spawning migrations, and potential spatial
770 overlap, but limited straying, among spawning populations throughout the year; Porch et al.,
771 2001; Goethel and Berger, 2017). Natal homing models need to account for relatively more
772 complex dynamics and may require additional data (e.g., natal origin of catch and surveys when
773 populations overlap during fishing seasons), which has limited their application (Li et al., 2015;
774 2018; Vincent et al., 2017). Thus, it is likely that incorrect assumptions about natal homing in a
775 spatial IPM (i.e., assuming it is occurring when it is not or ignoring it when it does occur) may
776 lead to large estimation bias and has been shown to lead to different interpretation of sustainable
777 harvest levels (Francis and McKenzie, 2015; Goethel and Berger, 2017).
778 The tradeoff between parsimony and complexity is a recurring issue within all types of
779 spatial models, especially regarding assumptions and parameterizations of population structure,
780 movement, recruit apportionment, and the number of spatial units to model. As model flexibility
781 and complexity increases, models are better able to emulate real world dynamics and reduce bias,
782 but there is a limit to the added complexity that can be adequately estimated in a spatial IPM,
783 especially as the number of units modeled increases and sample sizes decrease (Cope and Punt,
784 2011; Punt, 2019b). We demonstrated that with the *SH* IPM, estimation of recruit apportionment
785 can be problematic. Goethel et al. (2021) suggests using flexible, but adequately constrained
786 movement parameterizations, and that theory can likely be applied to other parameters (e.g.,
787 apportionment), while Punt (2019b) further supports parameter sharing across areas where such
788 an approach might be logical. Spatiotemporal IPMs (e.g., Cao et al., 2020), as opposed to the
789 spatially stratified approaches explored here, also demonstrate promise for reducing the number
790 of parameters and maximizing information content from observed data by directly accounting for
791 spatial correlation among fine-scale units. Future work to meld these two spatial IPM approaches
792 could help identify more robust methods to support spatial fisheries management.
793 Relatively little is known about the influence of spatial dynamics on levels of sustainable
794 harvest. Bosley et al. (2019) demonstrated that when movement was present, a broad range of
795 harvest rate combinations across areas led to maximum yield from the system. Similarly, Goethel
796 and Berger (2017) demonstrated that sustainable yield varied substantially depending on the
797 assumed population structure, movement patterns and rates, and the distribution of effort. Thus,

better understanding of spatial dynamics may lead to a rethinking of how target and limit biological reference points are developed and applied. Dynamic, time-varying connectivity across space and population components impedes the ability to achieve any single equilibrium rebuilding target and essentially spreads the impact of fishing across the entire spatial domain. Accounting for the multiscalar nature of spatial dynamics (e.g., local and regional interactions within and across metapopulation components) may not be fully tractable within the current reference point paradigm. Further development of spatial OMs that can be used to test alternate harvest control rules that account for desired spatial utilization of the resource, as well as the spatial dynamics of the species is required to determine truly sustainable management regimes.

5 Conclusions

Our results provide further evidence that spatial IPMs are generally robust to the diversity of spatial dynamics observed for marine resources and should be more widely applied when spatial structure is suspected. It also contributes to the growing body of work to support development of the "next generation' of fishery stock assessments (Punt et al., 2020). In the absence of knowledge on underlying population structure, assumptions of spatial heterogeneity or metapopulation structure within spatial IPMs are likely to provide relatively unbiased parameter estimates in most situations. However, it is important to maintain flexible parameterization of movement dynamics or the risk of parameter bias may be similar to ignoring spatial structure altogether. *Panmictic* IPMs may be able to accurately estimate system level population trends but rely on potentially poor performing catch allocation methods to assign quota to management sub-units when spatial management is required (Bosley et al., 2019). Fleets-as-areas models provide limited benefit and can be highly biased, suggesting that spatial IPMs or individual closed population models that match the scale of important population units (when data are limited to inform movement dynamics) should be preferred over FAA approaches. As the performance of spatial IPMs continues to be explored and better understood, we believe that the management of harvested natural resources will benefit from the increased application of spatially explicit modeling approaches.

Acknowledgements

We would like to thank the reviewers who provided insightful comments that improved the manuscript. We would also like to thank NOAA's National Marine Fisheries Service Office of Science and Technology branch for providing funding under the Stock Assessment Analytical

Methods (SAAM) project to support this research and the National Research Council Research Associate Program (NRC- RAP) for support through a postdoctoral fellowship.

Data Availability Statement

No empirical data were collected for this study. Data used for this study were produced through simulation. Novel model code for all simulations is available at https://github.com/KateBoz/Spatial_IPM.

References

Aguilée, R., Lamber, A., & Claessen, D. (2011). Ecological speciation in dynamic landscapes. *Journal of Evolutionary Biology,* 24, 2663-2677. https://doi.org/10.1111/j.1420-9101.2011.02392.x

Akçakaya, H. R., Radeloff, V. C., Mladenoff, D. J., & He, H. S. (2004). Integrating landscape and metapopulation modeling approaches: viability of the sharp-tailed grouse in a dynamic landscape. *Conservation Biology,* 18, 526-537. https://doi.org/10.1111/j.1523-1739.2004.00520.x.

Allen, A. M., & Singh, N. J. (2016). Linking movement ecology with wildlife management and conservation. *Frontiers in Ecology and Evolution*, **3**, 155. https://doi.org/10.3389/fevo.2015.00155.

Baranov, F. I. (1918). K voprosu o biologicheskikh osnovaniyakh rylmogo khozyaistva. [On the question of the biological basis of fisheries.] *Nauchn. Issled. Ikhtiologicheskii Inst. Izv.* 1, 81-128.

Berger, A. M., Goethel, D. R., Lynch, P. D., Quinn II, T., Mormede, S., McKenzie, J., & Dunn, A. (2017). Space oddity: the mission for spatial integration. *Canadian Journal of Fisheries and Aquatic Sciences,* 74, 1698-1716. https://doi.org/10.1093/icesjms/fsab096

Berger, A. M, Deroba, J. J., Bosley, K. M., Goethel, D. R., Langseth, B. J., Schueller, A. M., & Hanselman, D. H. (2020). Incoherent dimensionality in fisheries management: consequences of misaligned stock assessment and population boundaries. *ICES Journal of Marine Science*. https://doi.org/10.1093/icesjms/fsaa203.

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

Bravington, M. V., Skaug, H. J., & Anderson, E. C. (2016). Close-kin mark-recapture. S*tatistical Science,* **31**, 259-274. https://doi.org/10.1214/16-STS552

Brownie, C., Hines, J. E., Nichols, J. D., Pollock, K. H., & Hestbeck, J. B. (1993). Capture-recapture studies for multiple strata including non-Markovian transition probabilities. *Biometrics,* **49**, 1173-1187. https://doi.org/10.2307/2532259

Cadrin, S. X. (2020). Defining spatial structure for fishery stock assessment. *Fisheries Research,* **221**. https://doi.org/10.1016/j.fishres.2019.105397.

Cadrin, S. X., Kerr, L. A., & Mariani, S. (2014). Stock identification methods: an overview. In Cadrin, S. X., Kerr, L. A., & Mariani, S (Eds.),.*Stock identification methods: applications in fishery science*, 2nd Edition. Burlington, MA: Elsevier Science and Technology,. pp. 1-5.

Cadrin, S. X., Goethel, D. R., Morse, M. R., Gay, F., & Kerr, L. A. (2019). "So, where do you come from?" The impact of assumed spatial population structure on estimates of recruitment. *Fisheries Research,* 217, 156-168. https://doi.org/10.1016/j.fishres.2018.11.030.

Cao, J., Thorson, J. T., Punt, A., & Szuwalski, C. (2020). A novel spatiotemporal stock assessment framework to better address fine-scale species distributions: Development and simulation testing. *Fish and Fisheries,* 21, 350-367. https://doi.org/10.1111/faf.12433

Chandler, R. B., & Clark, J. D. (2014). Spatially explicit integrated population models. *Methods in Ecology and Evolution,* 5, 1351-1360. https://doi.org/10.1111/2041-210X.12153

Ciannelli, L., Fisher, J. A. D., Skern-Mauritzen, M., Hunsicker, M. E., Hidalgo, M., Frank, K. T., & Bailey, K. M. (2013). Theory, consequences and evidence of eroding population structure in harvested marine fishes: a review. *Marine Ecology Progress Series,* 480, 227-243. https://doi.org/10.3354/meps10067

Cope, J. M., & Punt, A. E. (2011). Reconciling stock assessment and management scales under conditions of spatially varying catch histories. *Fisheries Research,* 107, 22–38.

DeAngelis, D. L., & Yurek, S. (2017). Spatially explicit modeling in ecology: a review. *Ecosystems,* 20, 284-300. https://doi.org/10.1007/s10021-016-0066-z

de Moor, C. L., & Butterworth, D. S. (2015). Assessing the South African sardine resource: two stocks rather than one? *African Journal of Marine Science,* 37, 41-51. https://doi.org/10.2989/1814232X.2015.1009166

Denson, L. S., Sampson, D. B., & Stephens, A. (2017). Data needs and spatial structure considerations in stock assessments with regional differences in recruitment and exploitation. *Canadian Journal of Fisheries and Aquatic Sciences,* 74, 1918-1929. https://doi.org/10.1139/cjfas-2016-0277

Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across time and space. *Annual Reviews of Ecology, Evolution, and Systematics,* 40, 677-697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A., & Sibert, J. (2012). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software,* 27, 233-249. https://doi.org/10.1080/10556788.2011.597854

Francis, R. I. C. C., & McKenzie, J. R. (2015). Assessment of the SNA1 stocks in 2013. New Zealand Fisheries Assessment Report 2015/76. 82p.

Fraser, K. C., Davies, K. T. A., Davy, C. M., Ford, A. T., Flockhart, D. T. R., & Martins, E. G. (2018). Tracking the conservation promise of movement ecology. *Frontiers in Ecology and Evolution,* 6, https://doi.org/10.3389/fevo.2018.00150.

Fu, C., & Fanning, L. P. (2004). Spatial considerations in the management of Atlantic cod off Nova Scotia, Canada. *North American Journal of Fisheries Management,* 24, 775-784. https://doi.org/10.1577/M03-134.1

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

Goethel, D. R., Quinn II, T. R., & Cadrin, S. X. (2011). Incorporating spatial structure in stock assessment: movement modelling in marine fish population dynamics. *Reviews in Fisheries Science and Aquaculture,* 19, 119–136. https://doi.org/10.1080/10641262.2011.557451

Goethel, D. R., Legault, C. M., & Cadrin, S. X. (2015a). Demonstration of a spatially explicit, tag-integrated stock assessment model with application to three interconnected stocks of

yellowtail flounder off of New England. *ICES Journal of Marine Science,* **72**, 582–601. https://doi.org/10.1093/icesjms/fsu014

Goethel, D. R., Legault, C. M., & Cadrin, S. X. (2015b). Testing the performance of a spatially explicit tag-integrated stock assessment model of yellowtail flounder (*Limanda ferruginea*) through simulation analysis. *Canadian Journal of Fisheries and Aquatic Sciences,* **72**, 164–177. https://doi.org/10.1139/cjfas-2014-0244

Goethel, D. R., Bosley, K. M., Hanselman, D. H., Berger, A. M., Deroba, J. J., Langseth, B. J., & Schueller, A. M. (2019). Exploring the utility of different tag-recovery experimental designs for use in spatially explicit, tag-integrated stock assessment models. *Fisheries Research,* **219**. https://doi.org/10.1016.j.fishres.2019.105320.

Goethel, D. R., Bosley, K. M., Langseth, B. J., Deroba, J. J., Berger, A. M., Hanselman, D. H., & Schueller, A. M. (2021). Where do you think you're going? Accounting for ontogenetic and climate-induced movement in spatially stratified integrated population assessment models. *Fish and Fisheries,* **22**, 141-160. https://doi.org/10.1111/faf.12510.

Gu, W., Heikkila, R., & Hanski, I. (2002). Estimating the consequences of habitat fragmentation on extinction risk in dynamic landscapes. *Landscape Ecology,* **17**, 699-710. https://doi.org/10.1023/A:1022993317717

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

Hanselman, D. H, Heifetz, J., Echave, K. B., & Dressel, S. C. (2015). Move it or lose it: movement and mortality of sablefish tagged in Alaska. *Canadian Journal of Fisheries and Aquatic Sciences,* **72**, 238-251. https://doi.org/10.1139/cjfas-2014-0251

Hanselman, D. H., Rodgveller, C. J., Fenske, K. H., Shotwell, S. K., Echave, K. B., Malecha, P. W., & Lunsford, C. R. (2018). Assessment of the Sablefish Stock in Alaska. NPFMC Bering Sea, Aleutian Islands and Gulf of Alaska SAFE report.

Hastings, A. (1990). Spatial heterogeneity and ecological models. *Ecology,* **2**, 426-428. https://doi.org/10.2307/1940296

Hidalgo, M., Secor, D. H., & Browman, H. I. (2016). Observing and managing seascapes: linking synoptic oceanography, ecological processes, and geospatial modelling. *ICES Journal of Marine Science,* 73, 1825-1830. https://doi.org/10.1093/icesjms/fsw079

951 Hulson, P-J. F., Miller, S. E., Ianelli, J. N., & Quinn II, T. J. (2011). Including mark-recapture
952 data into a spatial age-structured model: walleye pollock (*Theragra chalcogramma*) in the
953 eastern Bering Sea. *Canadian Journal of Fisheries and Aquatic Sciences,* 68, 1625–1634.
954 https://doi.org/10.1139/f2011-060
955 Hulson, P-J. F., Quinn II, T. J., Hanselman, D. H., & Ianelli, J. N. (2013). Spatial modelling of
956 Bering Sea walleye Pollock with integrated age-structured assessment models in a changing
957 environment. *Canadian Journal of Fisheries and Aquatic Sciences,* 70, 1-15.
958 https://doi.org/10.1139/f2011-060
959 Hurtado-Ferro, F., Punt, A. E., & Hill, K. T. (2014). Use of multiple selectivity patterns as a
960 proxy for spatial structure. *Fisheries Research*, 158, 102-115.
961 https://doi.org/10.1016/j.fishres.2013.10.001
962 Hutchinson, W. F. (2008). The dangers of ignoring stock complexity in fishery management: a
963 case of the North Sea cod. *Biology Letters,* **4**, 693-695.
964 https://doi.org/10.1098/rsbl.2008.0443
965 Kareiva, P. (1990). Population dynamics in spatially complex environments: theory and data.
966 *Philosophical Transactions of the Royal Society of London,* **330**, 175-190.
967 https://doi.org/10.1098/rstb.1990.0191
968 Kerr, L. A., Cadrin, S. X., & Kovach, A. I. (2014). Consequences of a mismatch between
969 biological and management units on our perception of Atlantic cod off New England. *ICES*
970 *Journal of Marine Science,* **71**, 1366–1381. https://doi.org/10.1093/icesjms/fsu113
971 Kerr, L. A., Hintzen, N. T., Cadrin, S. X., Clausen, L. W., Dickey-Collas, M., Goethel, D.
972 R., Hatfield, E. M. C., Kritzer, J. P., & Nash, R. D. M. (2017). Lessons learned from
973 practical approaches to reconcile mismatches between biological population
974 structure and stock units of marine fish. *ICES Journal of Marine Science,* **74**, 1708-
975 1722. https://doi.org/10.1093/icesjms/fsw188
976 Keymer, J. E., Marquet, P. A., Velasco-Hernández, J. X., & Levin, S. A. (2000). Extinction
977 thresholds and metapopulation persistence in dynamic landscapes. *The American*
978 *Naturalist,* **156**, 478-494. https://doi.org/10.1086/303407
979 Lee, H. -H., Piner, K., Maunder, M., Taylor, I., & Methot, R. (2017). Evaluation of alternative
980 modelling approaches to account for spatial effects due to age-based movement. *Canadian*

Journal of Fisheries and Aquatic Sciences, **74**, 1832-1844. https://doi.org/10.1139/cjfas-2016-0294

Li, Y., Bence, J. R., & Brenden, T. O. (2015). An evaluation of alternative assessment approaches for intermixing fish populations: a case study with Great Lakes lake whitefish. *ICES Journal of Marine Science,* **72**, 70-81. https://doi.org/10.1093/icesjms/fsu057

Li, Y., Bence, J. R., & Brenden, T. O. (2018). Can spawning origin information of catch or a recruitment penalty improve assessment performance for a spatially structured stock assessment model? *Canadian Journal of Fisheries and Aquatic Sciences,* **75**, 2136-2148. https://doi.org/10.1139/cjfas-2017-0523

Little, R., Punt, A. E., Tuck, G., & Mapstone, B. (2017). Exploring the effect of sampling, protogyny, and larval advection on stock estimates subject to no-take closures in a spatially complex coral reef fishery on the Great Barrier Reef, Australia. *Canadian Journal of Fisheries and Aquatic Sciences,* **74**, 1950-1959. https://doi.org/10.1139/cjfas-2016-0349

Lowerre-Barbieri, S. K., Kays, R., Thorson, J. T., & Wikelski, M. (2019). The ocean's movescape: fisheries management in the bio-logging decade (2018-2028). *ICES Journal of Marine Science,* **76**, 477-488. https://doi.org/10.1093/icesjms/fsy211

Maunder, M. N., & Punt, A. E. (2013). A review of integrated analysis in fisheries stock assessment. *Fisheries Research*, **142**, 61-74. https://doi.org/10.1016/j.fishres.2012.07.025

McGilliard, C. R., Punt, A. E., Methot, R. D., & Hilborn, R. (2015). Accounting for marine reserves using spatial stock assessments. *Canadian Journal of Fisheries and Aquatic Sciences,* **72**, 262-280. https://doi.org/10.1139/cjfas-2013-0364

Methot, R. D., & Wetzel, C. R. (2013). Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research,* **142**, 86-99. https://doi.org/10.1016/j.fishres.2012.10.012

Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., & Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. *PNAS,* **105**, 19052-19059. https://doi.org/10.1073/pnas.0800375105

Northfield, T. D., Barton, B. T., & Schmitz, O. J. (2017). A spatial theory for emergent multiple predator-prey interactions in food webs. *Ecology and Evolution,* **7**, 6935-6948. https://doi.org/10.1002/ece3.3250 Ogburn, M. B., Harrison, A. –L., Whoriskey, F. G., Cooke, S. J., Flemming, J. E. M., & Torres, L. G. (2017). Addressing challenges in the

application of animal movement ecology to aquatic conservation and management. *Frontiers in Marine Science,* **4**, 10.3389/fmars.2017.00070. https://doi.org/10.3389/fmars.2017.00070

Petitgas, P., Secor, D. H., McQuinn, I., Huse, G., & Lo, N. (2010). Stock collapses and their recovery: mechanisms that establish and maintain life-cycle closure in space and time. *ICES Journal of Marine Science,* **67**, 1841-1848. https://doi.org/10.1093/icesjms/fsq082

Pickett, S. T. A., & Cadenasso, M. L. (1995). Landscape ecology: spatial heterogeneity in ecological systems. *Science,* **269**, 331-334. https://doi.org/10.1126/science.269.5222.331

Porch, C., Turner, S. C., & Powers, J. E. (2001). Virtual population analyses of Atlantic bluefin tuna with alternative models of transatlantic migration: 1970-1997. *Collective Volume of Scientific Papers ICCAT,* **52**, 1022–1045.

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

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., Little, L. R., & Tuck, G. N. (2017). The effect of marine closures on a feedback control management strategy used in a spatially aggregated stock assessment: a case study based on pink ling in Australia. *Canadian Journal of Fisheries and Aquatic Sciences,* **74**, 1960-1973. https://doi.org/10.1139/cjfas-2016-0017

Punt, A. E., Okamoto, D. K., MacCall, A. D., Shelton, A. O., Armitage, D. R., Cleary, J. S., Davies, I. P., Dressel, S. C., Francis, T. B., Levin, P. S., Jones, R. R., Kitka, H., Lee, L. C., McIsaac, J. A., Poe, M. R., Reifenstuhl, S., Silver, J. J., Schmidt, J. O., Thornton, T. R., Voss, R., & Woodruff, J. (2018). When are estimates of spawning stock biomass for small pelagic fishes improved by taking spatial structure into account? *Fisheries Research*, **206**, 65-78. https://doi.org/10.1016/j.fishres.2018.04.017

1043 Punt, A. E. (2019a). Modelling recruitment in a spatial context: a review of current
1044 approaches, simulation evaluation of options, and suggestions for best practices.
1045 *Fisheries Research,* **217**, 140-155. https://doi.org/10.1016/j.fishres.2017.08.021.
1046 Punt, A. E. (2019b). Spatial stock assessment methods: a viewpoint on current issues and
1047 assumptions. *Fisheries Research,* **213**, 132-143.
1048 https://doi.org/10.1016/j.fishres.2019.01.014
1049 Punt, A. E., Dunn, A., Pór Elvarsson, B., Hampton, J., Hoyle, S. D., Maunder, M. N.,
1050 Methot, R. D., & Nielsen, A. (2020). Essential features of the next-generation
1051 integrated fisheries stock assessment package: A perspective. *Fisheries Research,*
1052 **229**, 105617. https://doi.org/10.1016/j.fishres.2020.105617
1053 Smedbol, R. K., & Stephenson, R. L. 2001. The importance of managing within-species
1054 diversity in cod and herring fisheries of the North-Western Atlantic. *Journal of Fish*
1055 *Biology,* **59**, 109–128. https://doi.org/10.1111/j.1095-8649.2001.tb01382.x
1056 Struve, J., Lorenzen, K., Blanchard, J., Börger, L., Bunnefeld, N., Edwards, C., Hortal, J.,
1057 MacCall, A., Matthiopoulos, J., Van Moorter, B., Ozgul, A., Royer, F., Singh, N.,
1058 Yesson, C., & Bernard, R. (2010). Lost in space? Searching for directions in the
1059 spatial modelling of individuals, populations and species ranges. *Biological Letters,*
1060 **6**, 575-578. https://doi.org/10.1098/rsbl.2010.0338
1061 Thorson, J. T., Hicks, A. C., & Methot, R. D. (2015). Random effect estimation of time-
1062 varying factors in Stock Synthesis. *ICES Journal of Marine Science,* **72**, 178-185.
1063 https://doi.org/10.1093/icesjms/fst211
1064 Trainor, A. M., Schmitz, O. J., Ivan, J. S., & Shenk, T. M. (2014). Enhancing species
1065 distribution modeling by characterizing predator-prey interactions. *Ecological*
1066 *Applications,* **24**, 204-216. https://doi.org/10.1890/13-0336.1
1067 Vincent, M. T., Brenden, T. O., & Bence, J. R. (2017). Simulation testing the robustness of a
1068 multi-region tag-integrated assessment model that exhibits natal homing and estimates
1069 natural mortality and reporting rate. *Canadian Journal of Fisheries and Aquatic Sciences,*
1070 **74**, 1930-1949. https://doi.org/10.1139/cjfas-2016-0297
1071 Waterhouse, L., Sampson, D. B., Maunder, M., & Semmens, B. X. (2014). Using areas-as-fleets
1072 selectivity to model spatial fishing: asymptotic curves are unlikely under equilibrium
1073 conditions. *Fisheries Research*, **158**, 15-25. https://doi.org/10.1016/j.fishres.2014.01.009

White, L. A., Forester, J. D., & Craft, M. E. (2018). Dynamic, spatial models of parasite transmission in wildlife: their structure, applications and remaining challenges. *Journal of Animal Ecology,* **87**, 559-580. https://doi.org/10.1111/1365-2656.12761

Whittington, J., & Sawaya, M. A. (2015). A comparison of grizzly bear demographic parameters estimated from non-spatial and spatial open population capture-recapture models. *PLoS One,* **10**, 1-17. https://doi.org/10.1371/journal.pone.0134446

Ying, Y., Chen, Y., Lin, L., & Gao, T. (2011). Risks of ignoring fish population spatial structure in fisheries management. *Canadian Journal of Fisheries and Aquatic Sciences,* **68**, 2101–2120. https://doi.org/10.1139/f2011-116

Zipkin, E. F., & Saunders, S. P. (2018). Synthesizing multiple data types for biological conservation using integrated population models. *Biological Conservation,* **217**, 240-250. https://doi.org/10.1016/j.biocon.2017.10.017

Table 1: Study design with scenario descriptions and estimation model convergence rates. SH: Spatial heterogeneity, Metapop: Metapopulation, FAA: Fleets-as-Areas. Scenarios are denoted by *IPM:OM:Movement:Tags*

Scenario Name	Integrated Population Model	Estimated Movement	Recruitment Apportionment in IPM	Operating Model	Recruitment Apportionment in OM	Convergence Rate
Panmictic:Uniform	Panmictic	None	-	Uniform	Fixed - equal	100%
Panmictic:SH		None	-	SH	Fixed - unequal	100%
Panmictic:Metapop		None	-	Metapop	-	100%
FAA:Uniform	Fleets-as-Areas	None	Fixed - equal	SH	Fixed - equal	100%
FAA:SH		None	Fixed - equal	SH	Fixed - unequal	100%
FAA:Metapop		None	Fixed - equal	Metapop	-	100%
SH:Uniform:NM	Spatial Heterogeneity	None	Fixed - unequal	Uniform	Fixed - equal	100%
SH:SH:NM		None	Fixed - unequal	SH	Fixed - unequal	100%
SH:SH:Const		Constant	Fixed - unequal	SH	Fixed - unequal	100%
SH:SH:Const:Tags		Constant	Fixed - unequal	SH w/tags	Fixed - unequal	100%
SH:SH:Age:Tags		Age-varying	Fixed - unequal	SH w/tags	Fixed - unequal	91%
SH:Metapop:NM		None	Fixed - unequal	Metapop	-	100%
SH:Metapop:Const		Constant	Fixed - unequal	Metapop	-	100%
SH:Metapop:Const:Tags		Constant	Fixed - unequal	Metapop w/tags	-	100%
SH:Metapop:Age:Tags		Age-varying	Fixed - unequal	Metapop w/tags	-	97%
Metapop:Uniform	Metapopulation	None	-	Uniform	Fixed - equal	100%
Metapop:SH:NM		None	-	SH	Fixed - unequal	100%
Metpop:SH:Const		Constant	-	SH	Fixed - unequal	97%
Metpop:SH:Const:Tags		Constant	-	SH w/tags	Fixed - unequal	100%
Metpop:SH:Age:Tags		Age-varying	-	SH w/tags	Fixed - unequal	95%
Metapop:Metapop:NM		None	-	Metapop	-	100%

Metapop:Metapop:Const		Constant	-	Metapop	-	100%
Metapop:Metapop:Const:Tags		Constant	-	Metapop w/tags	-	100%
Metapop:Metapop:Age:Tags		Age-varying	-	Metapop w/tags	-	99%

Table 2: Table of parameter values used in the operating models. Note that the same total average system-wide recruitment is used for all three OM configurations, but each model assumes slightly different recruitment dynamics by area. Recruitment is the number of recruits entering the system and mortality terms are instantaneous rates (yr^{-1}). Recruit apportionment is the proportion of system-wide average recruitment that is assigned to each area.

Area	Parameters	Uniform	Spatial Heterogeneity	Metapopulation
Entire Spatial Domain	Average System Recruitment (R_{ave})	15,543,790	15,543,790	15,543,790
	Natural Mortality (M)	0.1	0.1	0.1
Area 1	Minimum Fishing Mortality (F_{min})	0.05	0.02	0.02
	Maximum Fishing Mortality (F_{max})	0.40	0.40	0.40
	Average Area-Specific Recruitment (R_{ave})	--	--	6,838,830
	Recruit Apportionment (ξ)	0.33	0.44	--
Area 2	Minimum Fishing Mortality (F_{min})	0.05	0.05	0.05
	Maximum Fishing Mortality (F_{max})	0.40	0.40	0.40
	Average Area-Specific Recruitment (R_{ave})	--	--	4,662,840
	Recruit Apportionment (ξ)	0.33	0.30	--
Area 3	Minimum Fishing Mortality (F_{min})	0.05	0.05	0.05
	Maximum Fishing Mortality (F_{max})	0.40	0.40	0.40
	Average Area-Specific	--	--	4,041,130

	Recruitment (R_{ave})			
	Recruit Apportionment (ξ)	0.33	0.26	--

Table 3. Table of *cumulative sCAPE* values for each model scenario. Scenarios are denoted by *IPM:OM:Movement:Tags*. Parameters for the *Panmictic* and *FAA* IPMs show only the system level *sCAPE* except for *F* in the *FAA* model where area specific parameters are estimated. Cells are shaded to error represent overall error level for a given parameter within each IPM scenario (e.g. only *Panmictic* IPMs are compared to each other and only *Spatial Heterogeneity* (*SH*) models are compared to each other). Bold cells represent the scenario with minimum error level within an IPM parameterization.

Scenario	SSB	Recruitment	F	Movement
Panmictic:Uniform	**0.002**	**0.012**	**0.008**	-
Panmictic:SH	0.045	0.022	0.093	-
Panmictic:Metapop	0.023	0.018	0.119	-
FAA:Uniform	**0.003**	**0.017**	**0.036**	-
FAA:SH	0.191	0.049	1.124	-
FAA:Metapop	0.107	0.023	0.848	-
SH:Uniform	**0.008**	**0.041**	**0.040**	-
SH:SH:NM	1.271	0.156	1.751	-
SH:SH:Const	0.344	0.048	0.487	0.687
SH:SH:Const:Tags	0.357	0.054	0.360	0.150
SH:SH:Age:Tags	0.113	0.053	0.144	0.097
SH:Metapop:NM	1.646	1.608	1.418	-
SH:Metapop:Const	0.294	1.436	0.417	0.325
SH:Metapop:Const:Tags	0.301	1.433	0.309	0.144
SH:Metapop:Age:Tags	0.096	1.429	0.155	**0.095**
Metapop:Uniform	**0.019**	**0.047**	**0.047**	-
Metapop:SH:NM	0.770	0.416	1.328	-
Metapop:SH:Const	0.758	0.389	1.704	0.908
Metapop:SH:Const:Tags	0.381	1.210	0.448	0.693
Metapop:SH:Age:Tags	0.120	0.616	0.149	0.608
Metapop:Metapop:NM	0.527	1.496	0.915	-
Metapop:Metapop:Const	0.388	1.338	0.839	2.674
Metapop:Metapop:Const:Tags	0.396	1.808	0.514	0.348
Metapop:Metapop:Age:Tags	0.098	0.947	0.153	**0.118**

Figure 1. Schematic illustrating the study design and demonstrating how the operating models (OM) were paired with the different integrated population models. Different shades represent spatial variation in fishery and biological parameters across areas in the *Spatial Heterogeneity* (*SH*) spatial structure (vertical dark lines delineate areas) or subpopulations (spaces between areas denote different subpopulations with varying demographic characteristics) for the *Metapopulation* spatial structure. Arrows indicate whether movement was modeled between areas (for the operating models) or estimated (in the IPM). The fleets-as-areas IPM was modeled as a uniform population with different fishery selectivity curves estimated for each area (pseudo-areas are delineated by dashed lines).

Figure 2: Scaled Cumulative Absolute Percent Error (*sCAPE*) for each model scenario. The *sCAPE* metric is scaled to the maximum value across all the scenarios and all the areas for each parameter. Note that no system level estimates of fishing mortality are available for spatial models because it is not straightforward to aggregate area-specific estimates to a system level total when different selectivity estimates exist for each area. Similarly, movement is only estimated by area and *sCAPE* values for movement represent the residency rate (i.e., one minus the total emigration from that area). Scenarios are denoted by *IPM:OM:Movement:Tags*.

Figure 3: Relative percent difference (*RPD*) between true and estimated values for *F* and *SSB* in the terminal year. Medians are represented by the solid points with 25th and 75th quartiles demarked by the solid lines within each violin plot. Zero bias is demonstrated by the dashed line. Note that no system level estimates of fishing mortality are available for spatial models because it is not possible to aggregate area-specific estimates to a system level total when there are different selectivity estimates for each area. Scenarios are denoted by *IPM:OM:Movement:Tags*.

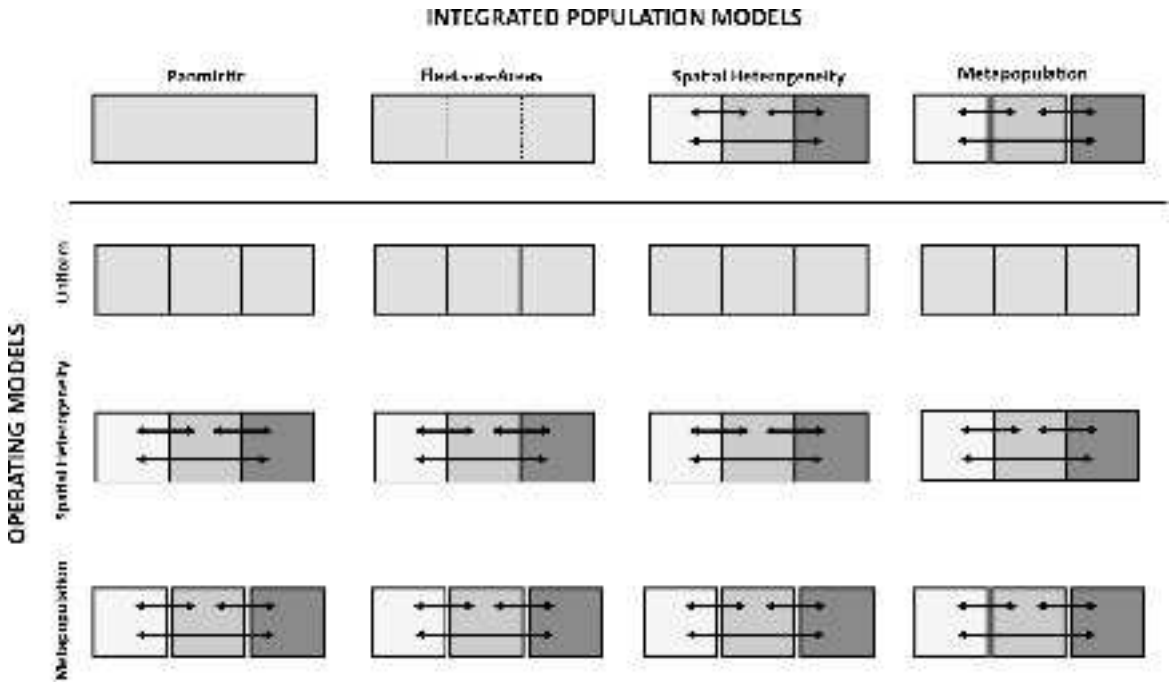
Figure 4: Relative percent difference (*RPD*) between true and estimated values of *SSB* for the panmictic and fleets-as-areas IPMs applied to all three operating models. Open points represent medians. Ribbons show the 100%, 90%, and 50% interquartile ranges. Zero bias is denoted by the dashed line. These IPMs are not spatially explicit, thus, no area-specific values are presented.

Figure 5: Relative percent difference (RPD) between true and estimated values of *SSB* for the *Spatial Heterogeneity* (*SH*) IPM with data from the *SH* OM. Open points represent medians. Ribbons show the 100%, 90%, and 50% interquartile ranges. Area is denoted by the right-hand panel titles. Zero bias is denoted by the dashed line.

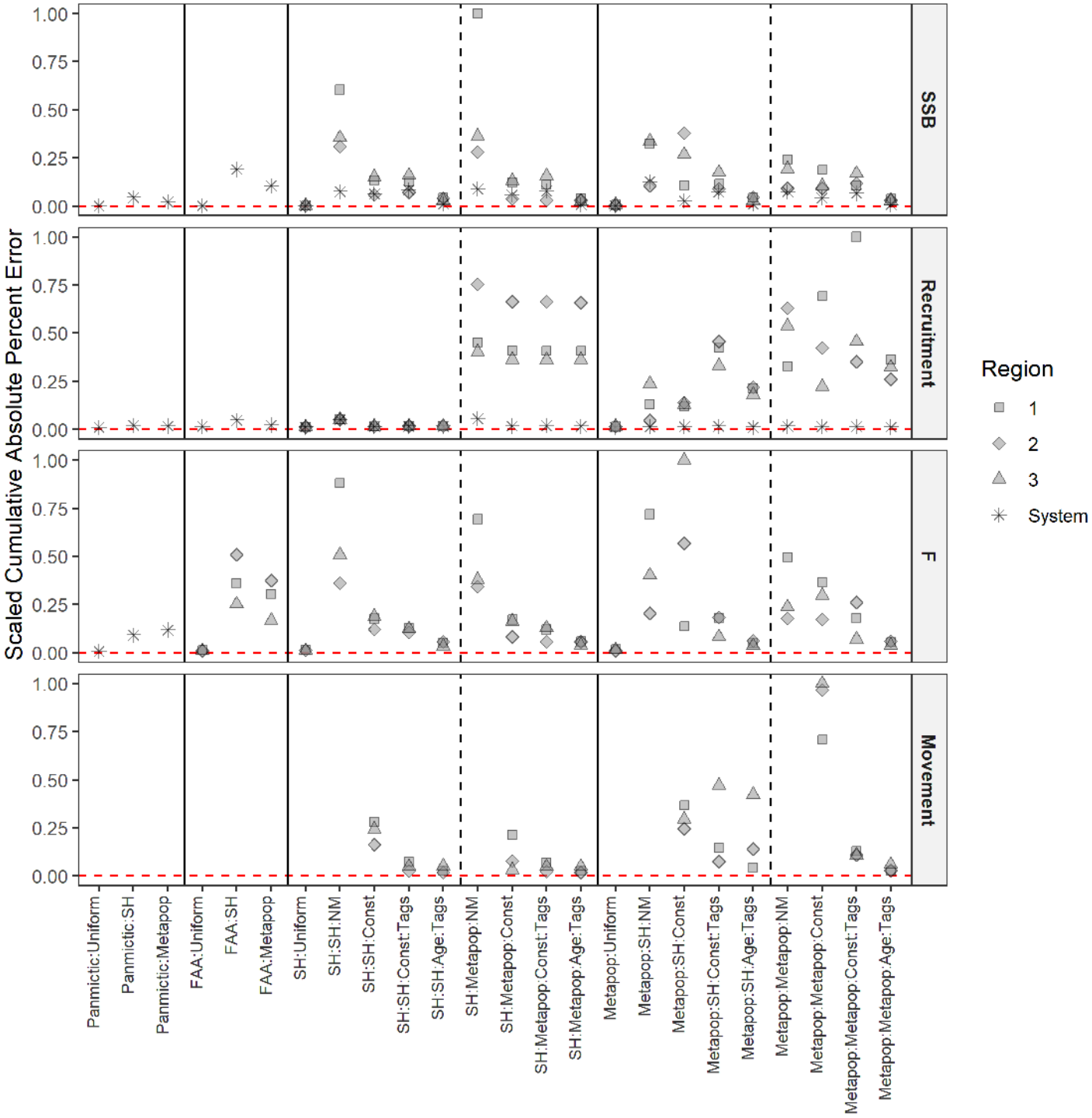
Figure 6: Relative percent difference (RPD) between true and estimated values of *SSB* for the *Spatial Heterogeneity* (*SH*) IPM with data from the *Metapopulation (Metapop)* OM with different parameterizations. Open points represent medians. Ribbons show the 100%, 90%, and 50% interquartile ranges. Area is denoted by the right-hand panel titles. Zero bias is denoted by the dashed line.

Figure 7: Relative percent difference (RPD) between true and estimated values for *SSB* for the *Metapopulation (Metapop)* IPM applied to the *Spatial Heterogeneity* (*SH)* OM with different parameterizations. Open points represent medians. Ribbons show the 100%, 90%, and 50% interquartile ranges. Area is denoted by the right-hand panel titles. Zero bias is denoted by the dashed line.

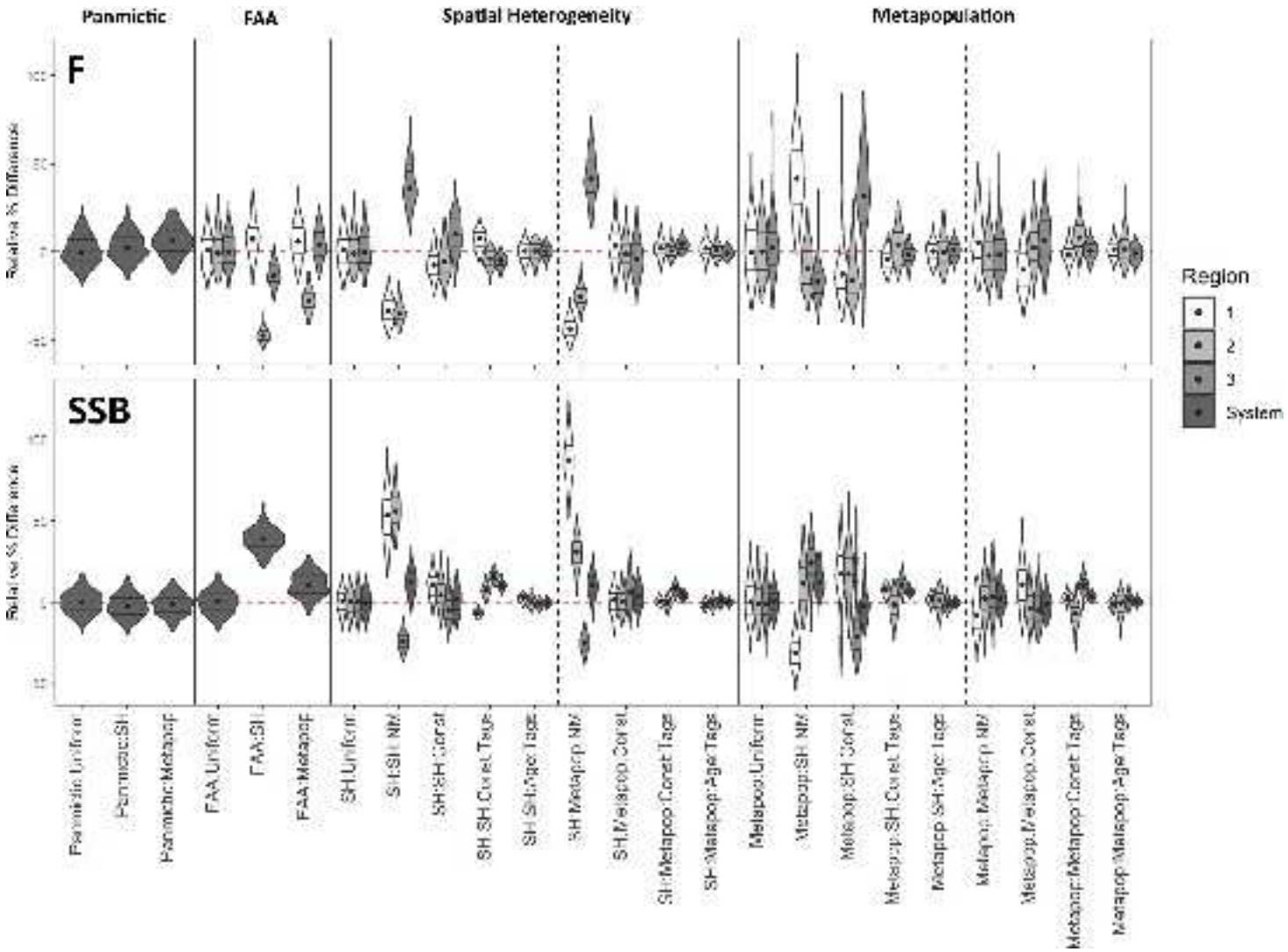
Figure 8: Relative percent difference (RPD) between true and estimated values for *SSB* for the *Metapopulation (Metapop)* IPM applied to the *Metapopulation* OM with different parameterizations. Open points represent medians. Ribbons show the 100%, 90%, and 50% interquartile ranges. Area is denoted by the right-hand panel titles. Zero bias is denoted by the dashed line.



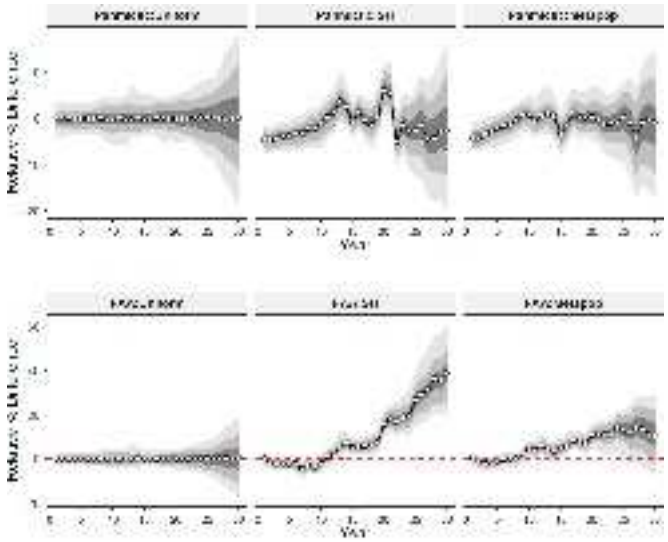
faf_12616_f1.png



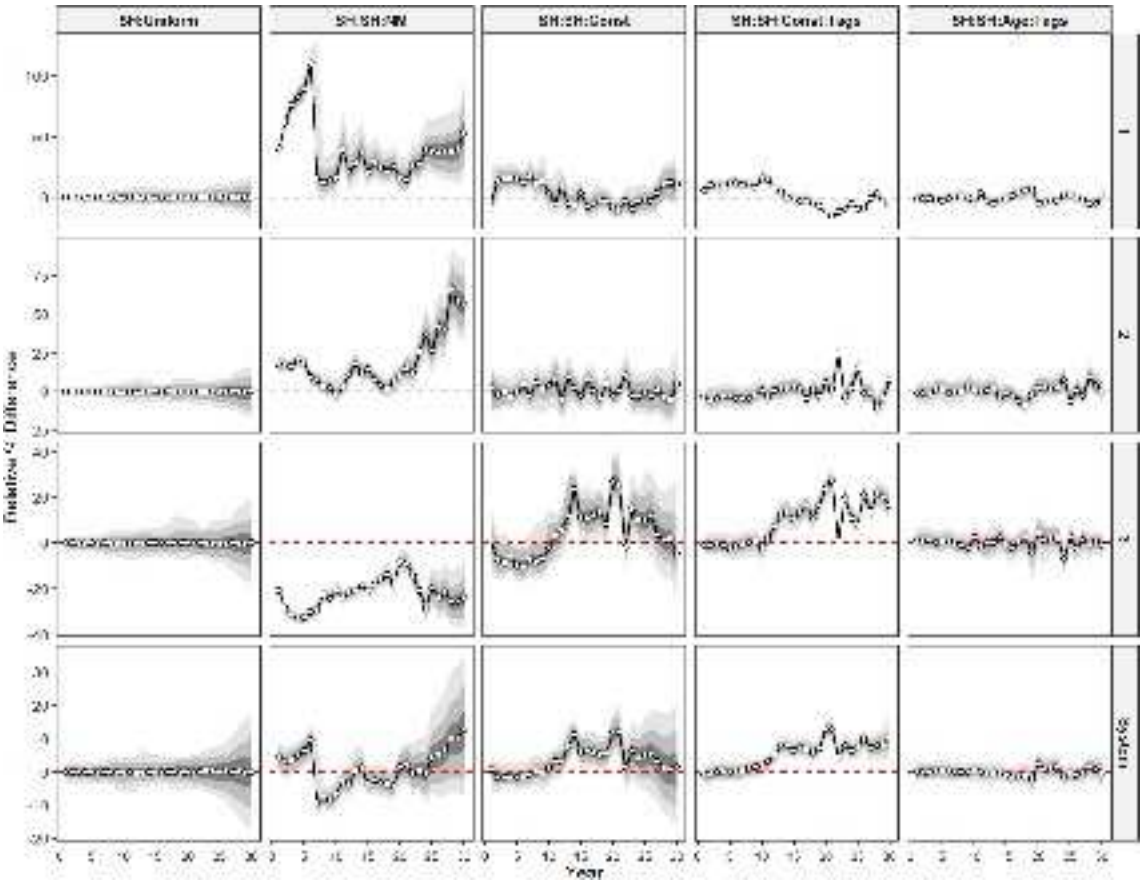
faf_12616_f2.png



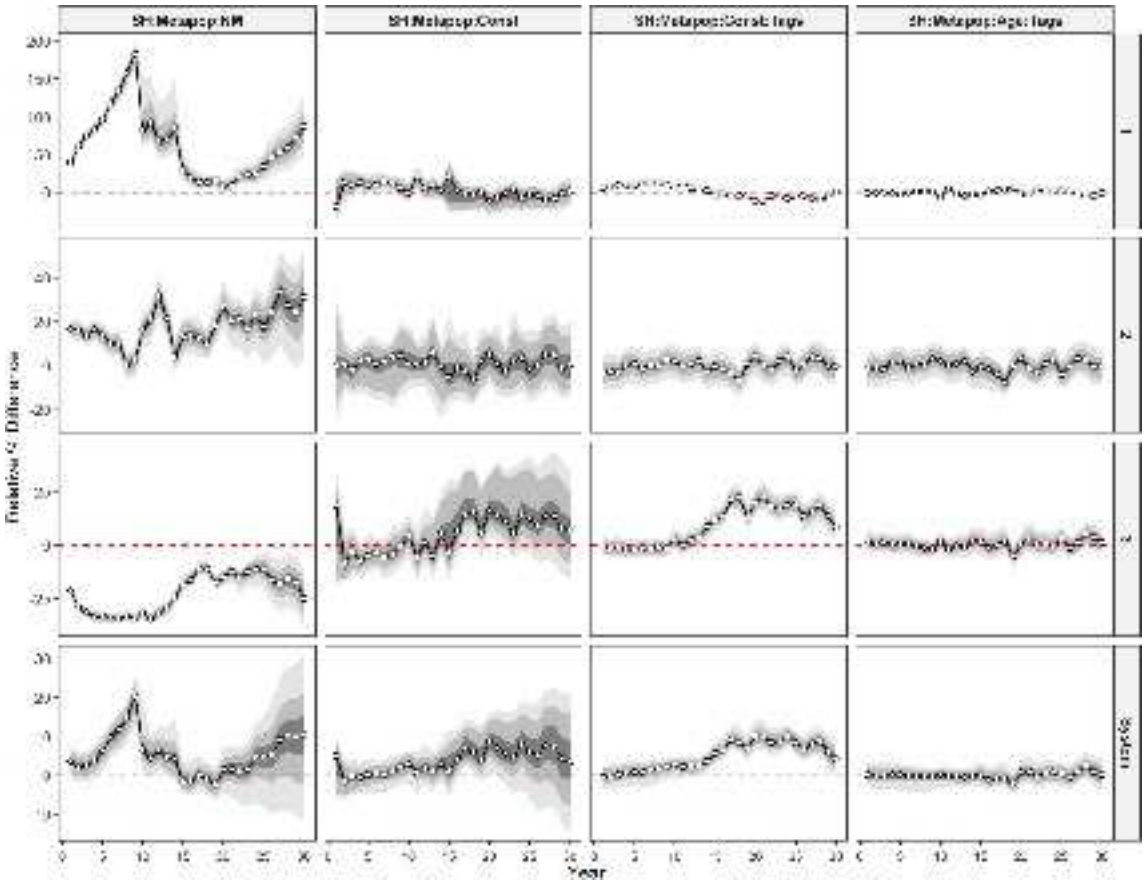
faf_12616_f3.png



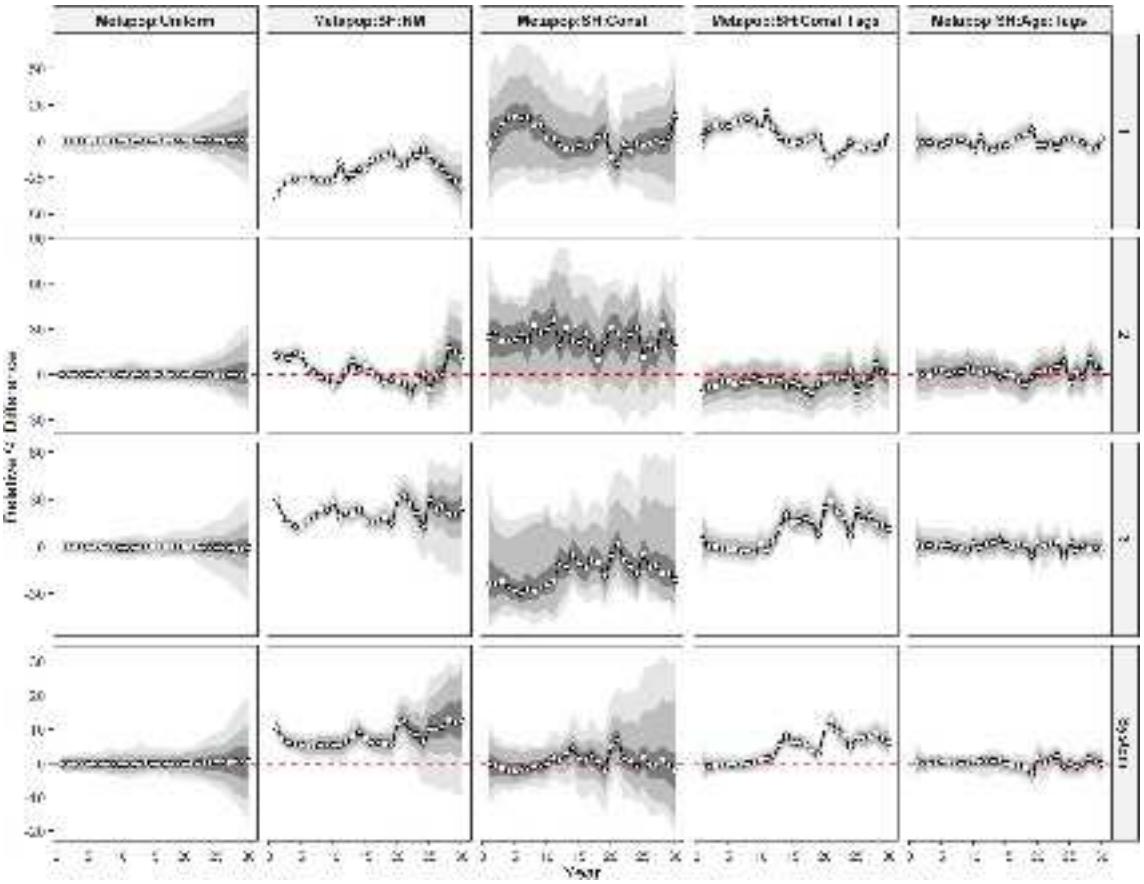
faf_12616_f4.png



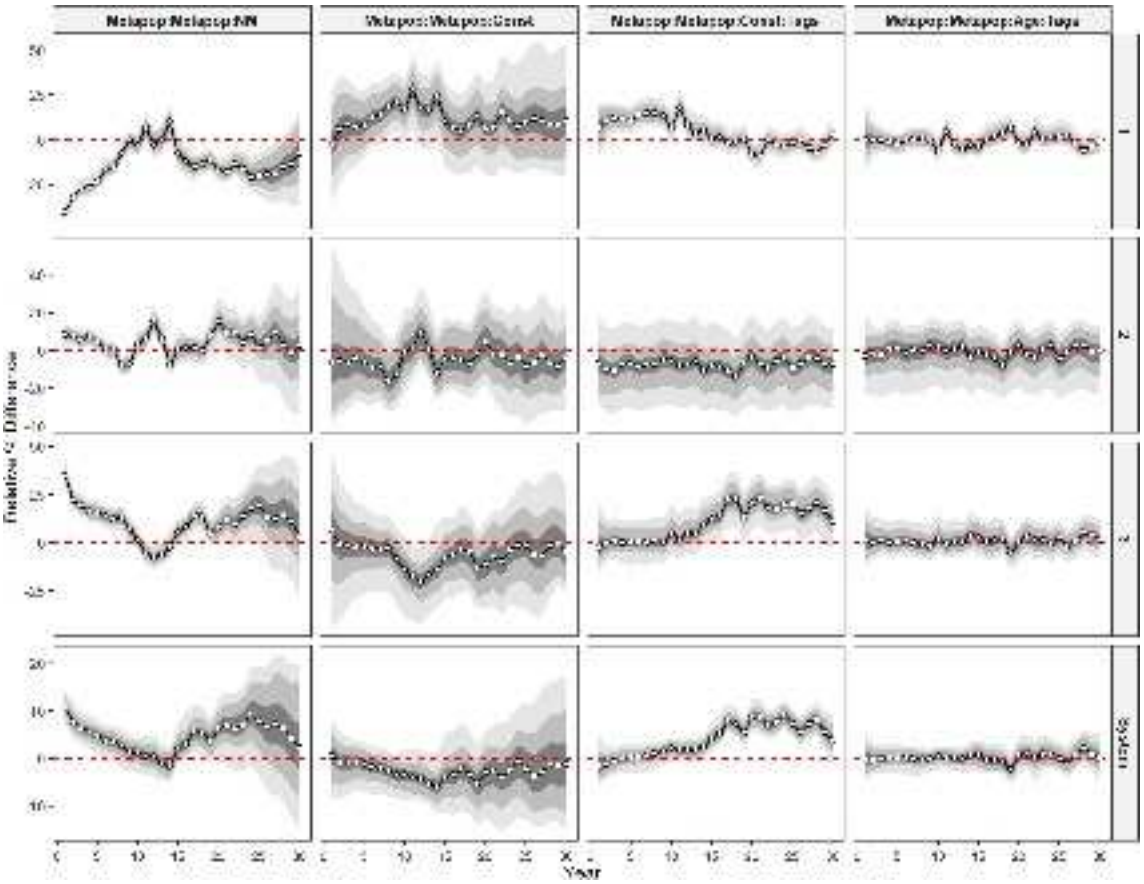
faf_12616_f5.png



faf_12616_f6.png



faf_12616_f7.png



faf_12616_f8.png