Canadian Journal of Fisheries and Aquatic Sciences
Journal canadien des sciences halieutiques et aquatiques

Effects of unequal capture probability on stock assessment abundance and mortality estimates: an example using the US Atlantic sea scallop fishery

Journal:	<i>Canadian Journal of Fisheries and Aquatic Sciences</i>
Manuscript ID	cjfas-2016-0296.R2
Manuscript Type:	Article
Date Submitted by the Author:	11-Mar-2017
Complete List of Authors:	Truesdell, Samuel; Michigan State University, Fisheries and Wildlife Hart, Deborah; National Marine Fisheries Service - NOAA Chen, Yong; University of Maine,
Please Select from this Special Issues list if applicable:	AFS Spatial Processes
Keyword:	Placopecten magellanicus, STOCK ASSESSMENT < General, catch-at-size, spatial fishing effort

SCHOLARONE™
Manuscripts

Effects of unequal capture probability on stock assessment abundance and mortality estimates: an example using the US Atlantic sea scallop fishery

Samuel B. Truesdell^{1,*} Deborah R. Hart² and Yong Chen¹

¹School of Marine Sciences, University of Maine, Orono, Maine 04469, USA

²Northeast Fisheries Science Center, 166 Water St., Woods Hole, MA 02543, USA

*Corresponding author (present address):

Samuel Truesdell
Quantitative Fisheries Center, MSU
375 Wilson Rd., UPLA Room 101
East Lansing, MI 48824
e-mail: truesd16@msu.edu
ph: 517-355-0126

Draft

Abstract

Most stock assessment models assume that the probability of capture for all individuals of the same size or age in the stock area is equal. However, this assumption is rarely, if ever, satisfied. We used spatially-referenced simulations, based on the US Atlantic sea scallop fishery, to generate catch, survey index, fishing effort and size structure data that we input into a (non-spatial) catch-at-size stock assessment model to estimate abundance and mortality rates. We show that spatial patterns in fishing mortality degrade model performance for sessile stocks. Fishing mortality tended to be overestimated and abundance underestimated because trends in fishing mortality were exaggerated and the model misestimated the numbers of larger individuals due to spatial fishing patterns. These results are particularly relevant to sedentary species such as scallops, but are applicable wherever strong spatial patterns exist in fishing mortality.

Draft

Introduction

Most stock assessment models are based on the assumption that each fish at a given size or age in a population has an equal probability of capture. However, the probability of capture in most fisheries is rarely equal over space. Orensanz et al. (2006) suggested potential explanations for this such as spatial patterns in the resource distribution, site attributes such as distance from port, and location-specific catchability. Even though spatial variability in fishing effort has long been recognized (e.g., Beverton and Holt 1957; Caddy 1975; Hart 2001), most stock assessments aggregate population and fishery processes over space. Models that assume that fishing mortality at a given size or age is uniform across the population are commonly used because they are simple (Orensanz and Jamieson 1998), mathematically practical to implement (Paloheimo and Dickie 1964), and because in many cases data are not sufficient to model spatially-referenced processes.

Recent efforts have been made to better integrate spatial components of populations and fisheries into standard stock assessment models (e.g., Goethel et al. 2011). The purpose is to improve assessments by aligning the spatial patterns in operating models more closely with real-world processes. Spatial processes such as dispersal (e.g., Morgan et al. 2000), migration (e.g., Ames 2004; Galuardi et al. 2010), growth (e.g., Brandt et al. 1992; McShane and Naylor 1995; Harris and Stokesbury 2006; Hart and Chute 2009), natural mortality rates (e.g., Merrill and Posgay 1964) and fishing mortality rates (e.g., Caddy 1975; Hart 2001; Aires-da-Silva et al. 2009) can all potentially affect catch-at-size (or catch-at-age) and population dynamics, and there has been some progress in integrating these spatial processes into models. For example, Hampton and Fournier (2001) developed a spatially disaggregated assessment model for yellowfin tuna that allowed the relationship between fish and fleet dynamics to vary over seven regions in the central and western Pacific. Hart et al. (2013) showed that running parallel assessments for areas open and closed to scallop fishing in the northeast US tended to outperform a

spatially aggregated version of the assessment model, and McGilliard et al. (2014) considered the more general question of the best way to assess a stock when a portion of it is protected by closed areas or “marine reserves.” Guan et al. (2013) used simulations to show that spatial heterogeneity can cause spatially aggregated catch-at-age models to underestimate biomass and overestimate fishing mortality. Truesdell et al. (2016) incorporated spatial variation in fishing effort and growth into spatially-averaged per-recruit analyses and demonstrated that spatial variability in these factors can substantially affect per-recruit curves and associated reference points.

This paper investigates the consequences of heterogeneous fishing mortality to stock assessment in the context of a sessile population with spatial patterns in fishing effort that is assessed and managed based on fishery-dependent and fishery-independent data. We loosely modeled our simulation after the US sea scallop (*Placopecten magellanicus*) fishery, which operates off the US east coast from the Gulf of Maine to the Mid-Atlantic and is one of the most valuable US fisheries, with 2013 ex-vessel landings worth \$467 million (NMFS 2013). Effort in this fishery is not distributed homogeneously (see Truesdell et al. 2016 Fig. 1 for an example), and is a function of the distribution of the scallop resource along with other socioeconomic factors such as distance from port, meat condition and regulations (especially rotational and long-term closed areas; Hart and Rago 2006). This fishery is a useful example because commercial-sized scallops are essentially sedentary (Caddy 1972; Hart and Chute 2004) and so do not redistribute themselves at the scale of the fishery (Orensanz et al. 2006); as such, heterogeneity in effort necessarily translates to spatial differences in fishing mortality (assuming spatially uniform catchability). Further, there is extensive survey data, including annual dredge surveys since 1979 (Hart and Rago 2006) and vessel monitoring systems since 1998 (Palmer and Wigley 2009), so there is auxiliary information on both the population and fleet dynamics.

In this study, we conduct simulations of the sea scallop fishery with varying degrees of heterogeneity in fishing effort. Data from these simulations are then employed as inputs to a catch-at-size (CAS) assessment model (e.g., Sullivan et al. 1990; Hart et al. 2013) in order to understand the effects of spatial patterns in fishing mortality on stock assessment model results.

Methods

Simulation

We simulated 100 years of a spatially-referenced scallop population, fishery, and a fishery-independent survey that we used to evaluate bias in a catch-at-size model. The simulated populations and fishery occurred on a grid where each grid component was 10 minutes latitude by 10 minutes longitude (Fig. 1). The R package PBSmapping (Schnute et al. 2015) was used to define a spatial grid and aggregate the data within units. The simulations were loosely based on the US sea scallop population and fishery in the Mid-Atlantic with some simplifications that enabled us to test our hypotheses in a straightforward manner. Parts of the simulation (e.g., catch) were aggregated over space and exported to the assessment model to serve as the data used for estimation. In each year of the simulation growth occurred first, instantaneously, followed by instantaneous recruitment. Fishing and natural mortality processes then acted throughout the year.

Growth

Growth was carried out via a growth transition matrix that was divided into 24 5 mm bins, beginning with the bin centered on 22.5 mm and ending with the bin centered on 142.5 mm. This matrix was identical to that used in the 2014 NMFS sea scallop stock assessment (NEFSC 2014) for the Mid-Atlantic, and was based on estimates of von Bertalanffy growth parameters and their variability among individuals from shell ring analysis (Hart and Chute 2009). The product of a numbers-at-size vector in year y , N_y , and the transition matrix (T) gives the numbers-at-size vector after growth:

$$N'_y = TN_y \quad (1)$$

94

95 where N'_y is the new numbers-at-size vector immediately following instantaneous growth. The growth
 96 matrix T was spatially invariant and does not account for mortality.

97 *Recruitment*

98 Simulated recruitment was divided into two components: overall recruit numbers and the spatial
 99 distribution of recruits. The components were independent random selections from temporal and
 100 spatial analyses of scallop data. The equation for annual recruitment at grid square g ($R_{y,g}$) was

$$R_{y,g} = \ddot{R}_y P_g^R \quad (2)$$

101

102 where $\ddot{R}_y$ is the annual overall recruitment in numbers. $\ddot{R}_y$ was a random selection of one year from
 103 the recruitment time series (1975-2009) as estimated by the National Marine Fisheries Service scallop
 104 stock assessment model (NEFSC 2010) for the Mid-Atlantic. P_g^R represents the proportion of recruits at
 105 grid square g . P^R , a vector of length G (the number of grid squares), was randomly selected from a set
 106 of proportional recruitment allocation vectors. This set of vectors was generated using spatial
 107 recruitment data from NMFS scallop surveys from 1982-2010. For each year in these data, each grid
 108 square g was assigned the arithmetic mean of the recruits from survey tows that occurred in that
 109 square in year y . Grid squares that were unsampled in year y were assigned a value of the 10th
 110 percentile of recruits sampled in that year. The number of recruits per grid square was then converted
 111 to a proportion of the overall total in year y .

New recruits were assumed to have grown into the first 10 size bins of the model (scallops measuring 22.5 – 67.5 mm) according to the proportions 0.04, 0.14, 0.14, 0.13, 0.13, 0.12, 0.11, 0.09, 0.07 and 0.04 as estimated in the 2014 NMFS stock assessment model (NEFSC 2014).

Fishery Selectivity

Fishery selectivity was modeled by a logistic function:

$$S_b = 1/(1 + e^{\alpha_0 - \alpha_1 H_b}) \tag{3}$$

where H_b is the shell height (height is the “size” in the catch-at-size model presented here) at the midpoint of bin b , and $\alpha_0=18.2$ and $\alpha_1=0.17$ are parameters based on the 2014 stock assessment as estimated in the 2005-2013 time block (NEFSC 2014).

Natural mortality

Natural mortality was fixed at 0.15, an assumed rate for scallops in the Mid-Atlantic that has been used in stock assessments (NEFSC 2010). This rate applied to all size classes and did not vary by year or in space.

Fishing mortality

We applied various scenarios involving differing levels of spatial heterogeneity and methods for assigning fishing mortality to locations to test the impact of this variability on stock assessment model results. Our scenarios ranged from spatially homogeneous to varying degrees of heterogeneity in fishing mortality. Each grid location in each year had a unique value for fully selected fishing mortality. The factors that defined this value were: (1) the overall annual trend in average fishing mortality; (2) the

degree of heterogeneity in fully selected fishing mortality; and (3) how each location-specific fully selected fishing mortality was assigned to each grid square.

Annual average fully selected fishing mortality refers to the stock-wide fishing mortality rate that would be represented in a spatially aggregated version of the Baranov catch equation (i.e., the fishing mortality that results in the aggregated catch). We varied annual average fishing mortality according to a temporal trend that was determined using simple linear interpolation between reference years. The reference years were 1, 28, 42, 55, 70, 82, 90 and 100. The respective fully selected fishing mortalities at each of these reference years were randomly selected from the uniform distribution $U(0.1, 0.9)$. The simulation was designed so that identical fishing mortality trends were used in each scenario (i.e., iteration 5 in all fishing mortality scenarios had the same trend).

We derived spatial heterogeneity in fishing mortality within a given year from vessel monitoring system (VMS) data from 1998 to 2011 (Palmer and Wigley 2009) using a multi-step process. VMS tracks the location of individual fishing vessels via satellite while they are at sea. These spatial data were pre-processed to include only estimated speeds between 2 and 5 knots: speeds under 2 knots were assumed to represent at-sea processing and speeds over 5 knots to represent steaming to or from a fishing location. The remaining records indicated approximate fishing effort. For each year, VMS data were projected onto the grid and each grid unit was assigned a value that was the mean of the VMS data located within that grid square for the year. Grids where no fishing occurred were assigned a value of zero. Notably, while the grids were used to generate plausible spatial distributions for heterogeneity in effort, the spatial effort values were not tied to their historic locations (determining the locations of effort in the model is discussed below).

These observed VMS-based effort distributions were altered depending on the heterogeneity scenario. The degree of heterogeneity in a scenario was governed by a parameter ν such that

$$E_y^* = E_y^\nu \tag{4}$$

where the vector E_y is the observed VMS effort and the vector E_y^* is the effort that has been adjusted for heterogeneity. The exponent ν is a non-negative number that affects the skew (i.e., heterogeneity) in effort. When $\nu < 1$, effort was less heterogeneous than the observed effort data and when $\nu > 1$ effort was more heterogeneous. Seven scenarios were tested (each with 50 iterations) where ν took values of 0, 0.25, 0.5, 0.75, 1.0, 1.25 and 1.5. When ν was 0 fishing mortality was homogeneous.

The next stage of determining fully selected fishing mortality was assigning each value of the vector E_y^* to a location. Two scenarios were applied for this assignment that affected the probability of capture for the individuals in space: (1) an “ideal free distribution” assumption (IDF; Fretwell and Lucas 1969; Gillis 2003), and (2) a “weighted ideal free distribution” (WIDF). For the IDF assignment, fully selected fishing mortality was related to available harvestable biomass by location, which was the sum of biomass-at-size multiplied by selectivity. The assignment matches the rank of the vector E_y^* with the rank of a vector of harvestable biomass-by-location. The location with the most harvestable biomass was assigned the highest fishing effort and the lowest harvestable biomass was assigned the lowest effort. The WIDF scenario used the same ranking approach, but the harvestable biomass was weighted by (the inverse of) the minimum of distances to five major fishing ports (Gloucester MA, New Bedford MA, Montauk NY, Cape May NJ and Virginia Beach VA), so more fishing effort occurred closer to shore than in the IDF scenarios (e.g., Caddy and Carocci 1999). This was meant to be a simple reflection of socioeconomic variables that drive fleet distribution such as regulations, fishing costs and bycatch

(Hilborn and Ledbetter 1979; Orensanz and Jamieson 1998; Holland and Sutinen 2000; Wilen *et al.* 2002).

The final step in the assignment of fishing mortality was to adjust the location-specific distributions of effort (E_y^*) so the stock-wide F matched the annual trend in fully-selected fishing mortality (the interpolated trend in fishing mortality discussed earlier, $\bar{F}_y$). This was accomplished by identifying a scalar, λ_y , that could be multiplied by the vector E_y^* to result in a final average fishing mortality (accounting for numbers-by-location and fishery selectivity) $\bar{F}_y$. The equation (based on the Baranov catch equation)

$$\sum_b \bar{F}_y S_b / (\bar{F}_y S_b + M) N (1 - e^{-\bar{F}_y S_b - M}) = \sum_g \sum_b \lambda_y E_{y,g}^* S_b / (\lambda_y E_{y,g}^* S_b + M) N_{y,b,g} (1 - e^{-\lambda_y E_{y,g}^* S_b - M}) \quad (5)$$

was solved numerically for λ_y in each year so that the catch estimated from the aggregated catch equation using the desired mean fishing mortality $\bar{F}_y$ (the left-hand side) was equal to the sum of the annual spatial catches (the right-hand side). Annual fully-selected fishing mortality by location was then.

$$F_{y,g} = \lambda_y E_{y,g}^* \quad (6)$$

Data for the assessment model

Annual fishing effort that was exported to the stock assessment model ($\tilde{E}_y$) was calculated using the relationship between fishing mortality and fishing effort:

189

$$\tilde{E}_y = \bar{F}_y/q_F \tag{7}$$

190 where q_F , the fishery catchability, was fixed at 0.001. Note that this effort differs from E_y^* and E_y^v
191 above which were used only to determine the relative spatial distribution of fishing mortality.

192 Simulation outputs for the assessment model that include observation error were total catch and
193 proportions-at-size from both the fishery catch and the abundance index. Total observed catch in
194 weight in year y , C_y^W , was calculated using the Baranov catch equation and known weight-at-size with a
195 lognormal error term:

196

$$C_y^W = \sum_b \bar{F}_{y,b}/(\bar{F}_{y,b} + M) N_{y,b} (1 - e^{-\bar{F}_{y,b} - M}) W_b e^{\varepsilon - (\sigma_C^2/2)} \tag{8}$$

197 where W_b is the weight-at-size for bin b , ε is a standard random normal number and σ_C^2 is the lognormal
198 error standard deviation for catch, set at 0.05 (a real-scale CV of approximately 0.05). The survey
199 abundance index was

200

$$I = q_I \sum_b A_b e^{\varepsilon - (\sigma_I^2/2)} \tag{9}$$

201 where q_I was the survey catchability (fixed at 0.001), A_b was the total abundance in size bin b and σ_I
202 was the standard deviation for the index, set at 0.05 (a real-scale CV of approximately 0.05). The survey
203 had a flat selectivity pattern, which is consistent with the NMFS scallop survey (NEFSC 2014) for scallops
204 over 40 mm. Our smallest size class had a lower bound of 20 mm, but we still assumed a flat selectivity
205 pattern for all scallops in order to simplify our analyses.

Observed proportions-at-size for the catch data and the abundance index were derived from sampling the multinomial distribution with an effective sample size of 75 where the probabilities were the proportions-at-size from the catch data or the population, respectively. An effective sample size of 75 was small relative to the 24 size bins in the assessment model, but reflected the uncertainty in composition often observed in fisheries data (Francis 2011; Truesdell et al. *in press*). The vector of observed catch proportions-at-size in year y , P_y , is

$$P_y = MN\left(C_y^N / \sum C_y^N, ESS\right) \quad (10)$$

where $MN()$ is a function returning a single realization from a multinomial distribution with effective sample size ESS and expected probabilities-at-size $C_y^N / \sum C_y^N$ where C_y^N is a vector of numbers-at-size in year y . The process was the same for obtaining composition samples from the abundance index, only the sampling was based on the population rather than the commercial catch (i.e., C in Eqn. 10 was replaced by N for population abundance). The P_y above did not have an index subscript for simplification.

Other information provided to the assessment model without error (not including processes that are assumed known such as natural mortality and weight-at-size) were the ratio of catch observation error standard deviation to (1) recruitment deviations standard deviation and (2) survey abundance index standard deviation so the relative scale for the errors was known. These were used to weight the recruitment and abundance index components of the likelihood function.

We repeated each scenario (i.e., IDF or WIDF crossed with each possible value for ν) 50 times so the analyses covered a range of variability over these iterations. The variability among simulations was a

combination of variability in absolute recruitment, the spatial distribution of recruits, how local population size interacted with spatial heterogeneity in fishing mortality for each scenario, and observation error for the catch and survey index and the proportions-at-size for the catch and survey.

Alternative Data Weighting

To gain additional context for our findings, we adjusted the data weights for both the commercial catch index and the commercial catch composition in the assessment model likelihood function. In alternative A1 the likelihood value for the fishery catch was decreased five-fold (i.e., that likelihood component was multiplied by 0.2) and in alternative A2 the effective sample size for the catch composition was decreased from 75 to 15.

Catch-at-size model

Population Dynamics

Some processes in the stock assessment model were linked directly to the simulation and not estimated. The initial scallop population numbers-at-size were taken directly from the simulation and were not estimated parameters. Growth in the assessment model was identical to that in the simulation (using the same growth transition matrix) and also involved no estimated parameters. It occurred instantaneously at the start of each year, as in the simulation. Natural mortality was fixed at 0.15 across all years and ages and was identical to the simulation natural mortality.

Recruitment occurred instantaneously at the start of the year directly after growth and involved a mixture of fixed and estimated parameters. The probability of recruits entering particular size classes was fixed at the actual values used in the simulation. The annual recruitment magnitudes were estimated parameters that arose from a random walk:

$$R_y = \begin{cases} e^{r_0} & \text{if } y = y_0 \\ e^{\log(R_{y-1}) + r_y^{dev}} & \text{if } y > y_0 \end{cases} \quad (11)$$

248

249 where R_y is recruitment in year y , r_0 is a log-scale estimate of recruitment in the first model year, r_y^{dev}
 250 is the estimated log-scale recruitment deviation in year y , and y_0 is the first model year.

251 Fishing mortality was assumed directly related to fishing effort through catchability:

$$\bar{F}_y = q_F \bar{E}_y \quad (12)$$

252

253 where fishing effort was input from the simulation and q_F was a parameter estimated on a log scale.

254 Fishing mortality during year y for size bin b , $F_{y,b}$, was the product of selectivity and fully selected
 255 fishing mortality:

$$F_{y,b} = S_b \bar{F}_y \quad (13)$$

256

257 where S_b is fishery selectivity in size bin b . Fishery selectivity was a logistic function (see Eqn. 3), where

258 α_0 and α_1 were parameters estimated on a log scale.

259 The survey index was assumed directly related to population abundance by the equation:

$$I = q_I \sum_b A_b \quad (14)$$

260

261 where q_I is the (log-scale estimated) survey index catchability and A_b is the total number of individuals
 262 in bin b . No selectivity pattern was estimated for the survey (i.e., selectivity was 1.0 for all size classes)
 263 which was consistent with the simulation.

264 Each of these processes described above occurred during annual time-steps. Population change over
 265 one time step can be written (similarly to Sullivan et al. 1990) as

$$N_{y+1,b'} = \sum_{b=1}^{b'} T_{b,b'} [N_{y,b} e^{-F_{y,b} - M_{y,b}}] + R_{y+1,b'} \quad (15)$$

266

267 where b' represents each size bin as it is filled according to the dynamic processes contributed from all
 268 bins $\leq b'$.

269 The model was fit using AD Model Builder (Fournier et al. 2012). The catch, survey index, and
 270 recruitment time series were fit assuming a normal likelihood for the log-scale catch/index estimates or
 271 for penalty parameters (i.e., recruitment deviations) and the proportions-at-size were fit assuming a
 272 multinomial likelihood. The negative log-likelihood function was

$$L = \left[n \log(\sigma_C) + \sum_y (\log(C_y^W) - \log(\hat{C}_y^W))^2 / 2\sigma_C^2 \right] \quad (16)$$

$$+ \left[n \log(\sigma_I) + \sum_y (\log(I_y) - \log(\hat{I}_y))^2 / 2\sigma_I^2 \right]$$

$$+ \left[n \log(\sigma_R) + \sum_y (r_y^{dev})^2 / 2\sigma_R^2 \right] - \left[ESS^C \sum_y p_y^C \log \hat{p}_y^C \right]$$

$$- \left[ESS^I \sum_y p_y^I \log \hat{p}_y^I \right]$$

273

where n is the number of years, σ_C is the standard deviation of the catch time-series residuals, σ_R is the standard deviation for the recruitment deviation series, σ_I is the standard deviation for the survey residuals, p_y^C and $\hat{p}_y^C$ are the observed and predicted catch proportions-at-size in the commercial catch, and p_y^I and $\hat{p}_y^I$ are the observed and predicted catch proportions-at-size in the survey catch. Predicted catch in weight, $\hat{C}_y^W$, was found using the Baranov catch equation (in the same manner as in the simulation) and the known weight-at-size.

The parameters σ_R and σ_I were determined using variance ratios while σ_C is an estimated parameter. These parameters are related through the fixed ratios

$$\sigma_R = \gamma_R \sigma_C \text{ and } \sigma_I = \gamma_I \sigma_C \quad (17)$$

where γ_R and γ_I are the ratios. γ_R and γ_I are exact quantities output directly from the simulation as the ratio of the sample standard deviation for log-scale catch observation error in the simulation and either the sample standard deviation of the annual recruitment deviations (r_y^{dev}) or the log-scale survey index observation error. This means that if σ_C was correctly estimated, σ_R and σ_I would also be correct. The median value across all iterations in all scenarios for γ_R was 20.8 and for γ_I was 0.93.

Years 1-60 (out of 100) were discarded from the simulations as a burn-in period so the assessment model covers years 61-100. Sixty years represents multiple generations of scallops (over 95% of an unfished population would die of natural mortality after twenty years based on the assumed $M = 0.15$), and maximum life span estimates for sea scallops range from 20 to 29 years (Serchuk et al. 1978; Naidu and Robert 2006).

293 The model parameters that were estimated were: the log of sigma for the commercial catch data
 294 ($\log(\sigma_C)$), the log of both logistic selectivity parameters ($\log(\alpha_0)$ and $\log(\alpha_1)$), the log of catchability for
 295 the fishery ($\log(q_F)$) and for the survey ($\log(q_I)$), the log of recruitment in the first year (r_0) and 39 log-
 296 scale random walk parameters for recruitment in each year after year 1 (r_y^{dev}). Each parameter was
 297 bounded by $[x - \delta(\text{abs}(x)), x + \delta(\text{abs}(x))]$ where x is the true (log scale) value from the simulation
 298 and $\text{abs}()$ is the absolute value function. δ was 0.5 for most parameters. δ for σ_C was 1.0; this made it
 299 more difficult for the model to estimate the overall scale of the variances for commercial and survey
 300 catch and for recruitment. The recruitment random walk vector was bounded by
 301 $[x - 0.5(\text{abs}(\max(x))), x + 0.5(\text{abs}(\max(x)))]$ so the bounds were based on the maximum true
 302 recruitment deviation. The model was given perfect starting values for fitting the data (e.g., the initial
 303 values for the recruitment deviations were the actual recruitment deviations). We took this approach
 304 because we were interested in issues that stemmed from unequal capture probability rather than
 305 whether or not the model became trapped at a local maximum of the likelihood surface. Thus our
 306 assessment model was very efficient at low values of v . Though we did not show the results here, we
 307 also fit the model with random starting values (between the minimum and maximum bounds for the
 308 parameters). Under those conditions, our conclusions did not change.

309 Comparisons of the assessment model estimates to the true simulation values were made for fishing
 310 mortality, catch, abundance of individuals > 100 mm and annual recruitment using root mean square
 311 error. Root mean square error for a variable x is

$$RMSE_x = \frac{\sqrt{[\sum_i \sum_y (x - \hat{x})^2]/n}}{\mu_x} \quad (18)$$

312

where i indicates iteration, y indicates year, n the total number of records over iterations and years and μ_x the mean of x .

Beverton-Holt estimator

Dynamic size-structured population models such as the CAS model presented here should outperform simpler models, given sufficient data. However, because of their complexity it is important to ensure that the results under these simulation conditions are generalizable to other assessment methods. We used the Beverton-Holt mortality estimator (Beverton and Holt 1956; Quinn and Deriso 1999) to compare against our results from the CAS model. This model requires only (mean) length and von Bertalanffy growth parameters to estimate total mortality. The estimator, assuming that recruitment, growth and selectivity are constant over time (and space), is

$$Z_y = K (L_\infty - \bar{L}_y) / (\bar{L}_y - L_c) \quad (19)$$

where Z_y is the total instantaneous mortality and $\bar{L}_y$ is the mean length in year y , K and L_∞ are von Bertalanffy growth parameters, and L_c is the initial length where individuals are considered fully vulnerable to the fishery. The assumptions of this model are that the growth parameters are known and have no individual variability, recruitment is constant, mortality is constant for all ages greater than the minimum selected size, and that the population has reached equilibrium. When this model was used to test the catch-at-size findings, these assumptions were respected in the simulation except that fishing mortality was allowed to vary spatially. Annual average fishing mortality was fixed at 0.2, recruitment was fixed at the mean recruitment over the time series and selectivity was 0 for scallops < 100 mm and 1 for scallops $\geq$ 100 mm (so L_c was set at 100 mm). The von Bertalanffy parameters were from Hart and Chute (2009), and were based on the same data used to develop the growth transition matrix. The

simulation was run for 200 years to ensure equilibrium conditions were met and the estimates were made using the final 100 years of data. The comparison to the simulation was made using fishing mortality which was

$$F_y = Z_y - M \tag{20}$$

(where M is the known natural mortality from the simulation).

Results

The scenarios varied dramatically in terms of the degree of heterogeneity in the fishing fleet and the behavior of the fishery with respect to the distribution of biomass (Fig. 1). While the IDF and WIDF simulations were generally similar in that fishing tended to occur primarily in areas with the most biomass, in the WIDF scenarios there was a discernable difference in how fishing mortality was distributed in that fishing mortality tended to be higher closer to the major ports that were represented (Fig. 1).

When ν , the level of heterogeneity, was low the CAS model was able to predict fishing mortality (Fig. 2), abundance (Fig. 3) and proportions-at-size (Fig. 4) reasonably well because there was little difference between the operating and assessment models; however at higher levels of ν in the IDF and WIDF scenarios, model accuracy was reduced (Table 1). Notably, there was a buildup of scallops in the plus group that was not accounted for in the CAS model (Fig. 4). The over-abundant plus group developed in lightly fished areas that experienced little directed effort. The plus group aside, there were generally more large individuals when spatial heterogeneity in fishing was low because the fleet did not direct its efforts towards locations with the highest biomass. When heterogeneity was high the fleet did target concentrations and there were generally fewer larger scallops (again, aside from in the plus group).

The general trend of the model overestimating fishing mortality and underestimating abundance in the IDF and WIDF scenarios was apparent when simulation summary statistics were viewed across all iterations (Fig. 5). There was not a dramatic difference in the performance of the assessment model under the IDF and WIDF scenarios, although in both cases while the WIDF could be quite biased it was slightly less so than under the IDF scenario (Fig. 5). The assessment model was well-behaved: the maximum gradient was less than 0.001 over 90% of the time (across all scenarios together) and all models produced Hessians.

The results of alternative scenario A1, where the landed catch data were down-weighted, were similar to the base scenario where the CAS model overestimated fishing mortality and underestimated numbers; however, the degree of abundance bias was slightly reduced in scenario A1 (Fig. 6). In scenario A2, where the commercial catch composition data were down-weighted (Fig. 7), the bias was not substantially reduced for fishing mortality (though it changed from generally positive bias in the baseline scenario to negative in A2) but bias was much lower for abundance. Interestingly, both abundance of individuals over 100 mm and fishing mortality tended to be biased low under these conditions.

The Beverton-Holt estimator resulted in the same directional bias as the CAS assessment (Fig. 8) for fishing mortality. The fishing mortality estimate varied between approximately 0.2 and 0.7.

Discussion

Most stock assessment models make the fundamental assumption of equal capture probability among individuals in the population of the same size or age. When this assumption was met, our assessment model was unbiased on average, or at least approximately so. However, this assumption is rarely, if ever, exactly satisfied by actual fisheries and our results indicate that violations of this assumption can

bias stock assessment results. As the disparity in capture probability among individuals in the population increased, the catch-at-size stock assessment model produced less accurate and less precise estimates of fishing mortality and harvestable numbers of individuals. For the IDF and WIDF scenarios at $\nu > 0$, fishing mortality was over-predicted on average by the model when spatial heterogeneity was high, while numbers were biased low. The direction of bias we found in the CAS model was also supported by our corresponding analysis using the Beverton Holt mortality estimator. As the direction of bias was consistent among the two approaches (both models over-predicted fishing mortality in the presence of heterogeneity in F), these findings may be applicable beyond simply integrated catch-at-age or catch-at-size models.

The WIDF model was expected to be less biased than the IDF model because the distance-to-shore weighting introduced some randomness in capture probability. This was meant to reflect the fact that fishing fleets do not follow an ideal free distribution and instead are impacted by factors such as regulations, fishing costs, bycatch, spatial management and lack of information regarding the spatial distribution of a stock (Hilborn and Ledbetter 1979; Orensanz and Jamieson 1998; Holland and Sutinen 2000; Wilen *et al.* 2002). The actual impact of the WIDF observed in our results, however, was minimal. As more socioeconomic rather than resource-dependent factors influence fleet distribution, the departure from IDF will become greater and less bias might be expected in the assessment model because capture probability is closer to random. If we had changed the degree of distance weighting we could have pushed the fishery closer to shore, forcing this scenario to deviate more from the IDF. We did not adjust the weights beyond the inverse of distance, because we did not have available information for realistic distance weights. Even though the WIDF scenario did not deviate much from the IDF under these conditions, we included it to emphasize that fisheries are driven by more than simply resource abundance and this can impact model performance. Thus the IDF scenario we present

is an extreme case and the actual impact of spatial heterogeneity in fishing mortality depends on the relative importance of abundance and other socioeconomic considerations (along with the quality of information about the resource distribution) that impact harvesters' decision-making.

When the simulations were run under assumptions of unequal capture probability, the stock assessment model provided less accurate estimates of fishing mortality and abundance. This is because assumptions about random sampling were not met (Paloheimo and Dickie 1964). Bias in the results was manifested through violations of catchability assumptions. Constant catchability is unlikely under conditions where there is substantial spatiotemporal variability in resource abundance and the fishery has no spatial restrictions and can target concentrations.

The phenomenon of unequal capture probability is not unique to the types of assessment models discussed here and has been highly developed in the mark-recapture literature (e.g., Carothers 1973; Jolly and Dickson 1982; Chao 1987). In such studies the recapture of individuals can allow for the estimation of variability in catchability and improved estimates of population size (Jolly and Dickson 1982). While these studies are not directly applicable to the types of stock assessment models described here (i.e., that do not involve capture-recapture data), they do offer information regarding the potential ranges of variability in catchability among individuals and the impact of catchability variation on population size estimates. This may be useful for bounding sensitivity analyses that seek to determine the potential impacts of unequal capture probability and would be especially relevant for mobile species assessments that must assume some spatial redistribution of the stock, unlike for the sedentary scallops discussed here where the spatial distribution of fishing effort is probably largely related to capture probability.

Spatiotemporal variability in catchability, which is directly related to variability in fishing mortality and fishing effort, is common (Salthaug and Aanes 2003; Wilberg et al. 2009), and the sources of this variability in time and space can be similar; they can be a function of changing fish behavior or harvester behavior/management. Including a time-varying component in catchability is common in catch-at-age and catch-at-size assessments, and annual catchability is often modeled as a random walk or white noise (e.g., Wilberg and Bence 2006). Real differences in catchability in time or space that are not captured in the assessment model (i.e., model misspecification) can bias the estimated relationship between abundance and a catch-per-unit effort index, potentially influencing model results. Assigning a spatial component to variability in catchability is less common than a temporal component, but can be implemented in some cases by strategies such as separating fisheries into fleets with different catchabilities. In addition, Salthaug and Aanes (2003) found annually estimated catchability was related to measures of fleet spatial concentration in two of four stocks they examined and suggested that an adjustment for fleet spatial dispersion could improve assessment estimates in some cases.

A major source of bias in this assessment model was the misrepresentation of size composition by the fishery-dependent data when effort was heterogeneous. Due to the directed nature of fishing under heterogeneous mortality, heavily fished areas developed truncated size distributions while lightly fished areas accumulated individuals in the plus group (see Fig. 4 and the individuals in the plus group not registered by the CAS model), changing the catch composition data and the effective selectivity (Sampson and Scott 2011). Thus the proportions-at-size in the catch were not representative of the population using the assumed fixed selectivity. On the other hand, the survey was representative because the probability of capture for scallops at all locations was equal so the survey size composition mirrored the population composition (save for sampling error governed by the survey effective sample size). The alternative weighting scenarios suggest that in this case the bias under IDF could be driven

more by the challenge of fitting the size composition data than by issues related to variability in overall landings. Under scenarios A1 and A2, when the weighting of the catch index or the catch composition data were reduced, the bias was reduced for the abundance data, especially under scenario A2 where the composition data were down-weighted (see Figs. 5-7). These results (along with the obvious misrepresentation by the CAS model of individuals in the plus group; Fig. 4) indicate that the composition data may be a primary source of bias (although the degree of down-weighting for the likelihood components was not evaluated for comparability). A thorough analysis of the impact of weighting on these results (i.e., testing a full range of weighting factors) would reveal a more complete relationship between the weighting of likelihood components and corresponding patterns in the bias of this assessment model. This could be an area of future research.

Domed selectivity (modeled, for example, with double logistic or gamma functions) can partially account for spatial patterns in fishing mortality (Caddy 1975; Sampson and Scott 2011; Hart et al. 2013). These allow for the accumulation of individuals in larger size classes that are not physically available to the fishery (or are simply fished at a lower rate), which can occur when fishing effort varies spatially. Stock assessments have used dome-shaped selectivities for this reason, including for Atlantic sea scallops (NEFSC 2010, 2014). It is possible that a dome-shaped selectivity model would improve the accuracy of this CAS model under spatial heterogeneity in fishing effort because bias (at least for abundance) appears driven at least in part by the commercial catch size composition. Domed selectivity could be considered as an extension to this model in future work.

The relatively low natural mortality used in the simulations was likely a contributing factor to observed biases. If individuals died at a higher rate, there would be less build-up in the plus-group and the bias in fishing mortality and numbers would likely be reduced. We treated natural mortality in the simulations as constant over time, space, and size class. Merrill and Posgay (1964) used data on the ratio of scallops

that die from natural causes to live scallops and gave evidence for variability in natural mortality over time and space, and there is evidence for size-specific natural mortality in the Mid-Atlantic (Shank et al. 2012). We did not consider these because our focus was on differences in fishing mortality, and simultaneously varying natural mortality would make our results harder to interpret. The Beverton-Holt equilibrium estimator for mortality was biased high when there was spatial heterogeneity in fishing effort, matching our results when we used the CAS model and making our conclusions more generalizable. Because of the non-linearity of the von Bertalanffy equation, a reduction in mortality affects the mean length less than an increase in mortality. Thus spatial heterogeneity in fishing effort reduces overall mean length and increases the estimated mortality from the Beverton-Holt estimator.

In this study, heterogeneity in effort was induced by socioeconomic choices of fishers to fish in areas of highest catch rates and in some cases in areas closer to port. Spatial heterogeneity can also be caused by explicit spatial management measures such as marine protected areas (MPAs) which would create different spatial effort patterns than the ones assumed here. Nonetheless, studies on the effects of MPAs on whole-stock assessments indicate that their effects are similar to those observed here, namely biomass tends to be underestimated and fishing mortality overestimated (Hart et al. 2013; McGilliard et al. 2014).

Similarly, Kerr et al. (2010) and Guan et al. (2013) found that aggregated stock models underestimated spawner biomass because of spatial variability in vital rates, recruitment and/or mortality. However, Caddy (1975) pointed out that using a “swept-area” approach to estimating fishing mortality (i.e., assuming that fishing mortality is proportional to effort) could underestimate F when effort is concentrated in areas of very high catch rates. Directional bias from this approach due to spatial heterogeneity is opposite from our CAS and Beverton-Holt conclusions. While our results were

consistent with much of the published literature, it is worth noting that the degree of bias and even the direction of bias depends on the assessment method that is employed.

The effects of spatial variability in fishing mortality on yield has also been investigated. In most cases, this variability decreases yield-per-recruit but increases biomass- or eggs-per-recruit (Hart 2001; Truesdell et al., 2016). At fishing mortalities around or less than F_{MSY} , this usually translates into reduced yield, because increased eggs-per-recruit typically does not increase recruitment substantially at high biomasses corresponding to relatively low fishing mortality rates (Hart 2006). However, under overfished conditions, increases in eggs-per-recruit may translate into sufficiently greater recruitment which more than compensates for the reduced yield per recruit. If inter-cohort density-dependence is dominant (i.e., adults suppress juvenile abundance), spatial heterogeneity in fishing mortality can increase yields even at high biomass levels (Ralston and O'Farrell 2008). However, intra-cohort density dependence is rarely dominant (Hart and Sissenwine 2009).

The results from this paper demonstrate that under these conditions fishing mortality can be overestimated when spatial heterogeneity in effort is ignored. This could superficially be considered a risk-averse result because the true fishing mortality is actually lower than the estimate. However, ignoring this issue and presenting biased assessment results for management decisions is inadvisable for two reasons. First, although average fishing mortality may be at what is considered a “safe” level for the stock, depending on the level of heterogeneity particular areas may experience considerably higher fishing mortality. If the assessment region does not comprise a unit stock, this could lead to serial depletion (e.g., Selgeby 1982; Ames 2004). Second, spatial variability in mortality would also affect fishery reference points (Hart 2001; Truesdell et al. 2016), so that comparisons of even the true fishing mortality rates with reference points computed under the assumption that fishing is spatially uniform may be inappropriate. Finally, assessment bias may also be an issue because it is possible that such

512 overestimated fishing mortality could lead to conservative management strategies that result in
513 foregone yield.

514 The degree of bias in these results may be overstated (relative to an actual stock assessment) because of
515 the data that are passed to the assessment model without error such as the growth transition matrix,
516 fishing effort and the likelihood standard deviation ratios. If these did include error we would expect to
517 see degraded performance of both the spatially homogeneous and spatially heterogeneous models and
518 possibly a reduced degree of directional bias due to increased noise. We chose to pass the these data to
519 the assessment model error-free to decrease the number of variables that included uncertainty so we
520 could focus on how heterogeneity was impacting the assessment model results.

521 Defining stock areas to homogenize spatial processes (Hart et al. 2013; McGilliard 2014) or integrating
522 space directly into assessment models (Hampton and Fournier 2001) are promising research avenues
523 that can lead to more accurate stock assessment. The challenge with respect to these methods,
524 however, is not the complicated model structure but the data requirements to estimate parameter
525 differences for fine-scale spatial processes. Spatially-integrated models may have a slight advantage
526 over independent models that cover different regions because of the potential to share parameters
527 among stock areas or to use hierarchical models, but data will still be an issue for many stocks.

528 This research examined the impact of heterogeneity in fishing mortality on assessment results for a
529 sedentary stock. It is a convenient case study because scallops do not distribute themselves when
530 vulnerable to the fishery so heterogeneity in fishing effort implies heterogeneity in fishing mortality.
531 Most stocks, however, are mobile so this assumption is not necessarily valid. Our study may represent
532 an extreme case of heterogeneity in effort (and an upper bound on resulting assessment model bias)
533 because of the lack of redistribution of the target species. The overall message emphasizes the

importance of understanding the interaction between variability in capture probability and the results of an assessment model; however, the disparity in capture probability and thus the impact on assessment will always be case-specific.

Most stock assessment models ignore spatial patterns in favor of spatially aggregated models. Thus, the types of spatial issues we examined here may contribute to similar biases in real-world assessments.

We recommend that simulation-based sensitivity analyses such as those presented here be undertaken to help inform decision-making in cases where it is not possible to implement spatial dynamics directly into a model or to administer separate assessments for areas that experience very different fishing mortalities.

Acknowledgements

The authors would like to thank Larry Jacobson and two anonymous reviewers for helpful comments on this manuscript. This research was funded by a NMFS-Sea Grant population dynamics fellowship and the University of Maine Correll Fellowship. This is publication 2017-09 of the Quantitative Fisheries Center at Michigan State University.

Sources cited

- Aires-da-Silva, A.M., Maunder, M.N., Gallucci, V.F., Kohler, N.E., and Hoey, J.J. 2009. A spatially structured tagging model to estimate movement and fishing mortality rates for the blue shark (*Prionace glauca*) in the North Atlantic Ocean. *Mar. Freshwater Res.* 60: 1029–1043.
- Ames, E.P. 2004. Atlantic cod stock structure in the Gulf of Maine. *Fisheries* 29: 10–28.
- Beverton, R.J.H. and S.J. Holt. 1956. A review of methods for estimating mortality rates in fish populations, with special reference to sources of bias in catch sampling. *Rapp. P.-v. Reun. Cons. Int. Explor. Mer.* 140:67-83.
- Beverton, R.J., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Chapman and Hall, London.
- Brandt, S.B., Mason, D.M., and Patrick, E.V. 1992. Spatially-explicit models of fish growth rate. *Fisheries* 17: 23–35.
- Caddy, J.F. 1972. Progressive loss of byssus attachment with size in the sea scallop, *Placopecten magellanicus* (Gmelin). *Journal of Experimental Marine Biology and Ecology* 9(2): 179–190.
- Caddy, J. 1975. Spatial model for an exploited shellfish population, and its application to Georges Bank scallop fishery. *Journal of the Fisheries Research Board of Canada* 32: 1305–1328.
- Caddy, J., and Carocci, F. 1999. The spatial allocation of fishing intensity by port-based inshore fleets: a GIS application. *ICES Journal of Marine Science* 56: 388–103.

- 566 Carothers, A.D. 1973. Capture-recapture methods applied to a population with known parameters. J.
567 Animal Ecol.: 125–146.
- 568 Chao, A. 1987. Estimating the population size for capture-recapture data with unequal catchability.
569 Biometrics 43: 783–791.
- 570 Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., Sibert, J.R.
571 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized
572 complex nonlinear models. Optim. Methods Softw. 27, 233-249.
- 573 Fretwell, S.D., and Lucas, H.L. 1969. On territorial behavior and other factors influencing habitat
574 distribution in birds. Acta Biotheor. 19(1): 16–36.
- 575 Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat.
576 Sci. 68: 1124-1138.
- 577 Galuardi, B., Royer, F., Golet, W., Logan, J., Neilson, J., and Lutcavage, M. 2010. Complex migration
578 routes of Atlantic bluefin tuna (*Thunnus thynnus*) question current population structure paradigm. Can.
579 J. Fish. Aquat. Sci. 67: 966–976.
- 580 Gillis, D.M. 2003. Ideal free distributions in fleet dynamics: a behavioral perspective on vessel movement
581 in fisheries analysis. Can. J. of Zoolog. 81: 177–187.
- 582 Goethel, D.R., Quinn, T.J., and Cadrin, S.X. 2011. Incorporating Spatial Structure in Stock Assessment:
583 Movement Modeling in Marine Fish Population Dynamics. Rev. Fish. Sci. 19: 119–136.

- 584 Guan, W., Cao, J., Chen, Y., and Cieri, M. 2013. Impacts of population and fishery spatial structures on
585 fishery stock assessment. *Can. J. Fish. Aquat. Sci.* 70: 1178–1189.
- 586 Hampton, J., and Fournier, D.A. 2001. A spatially disaggregated, length-based, age-structured population
587 model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. *Mar. Freshwater*
588 *Res.* 52: 937–963.
- 589 Hart, D.R., 2001. Individual-based yield-per-recruit analysis, with an application to the Atlantic sea
590 scallop, *Placopecten magellanicus*. *Can. J. Fish. Aquat. Sci.* 58(12):2351-2358.
- 591 Hart, D.R. 2006. When do marine reserves increase fishery yield? *Can. J. Fish. Aquat. Sci.* 63: 1445–1449.
- 592 Hart, D.R., and Chute, A.S. 2004. Sea scallop, *Placopecten magellanicus*, life history and habitat
593 characteristics. NOAA Technical Memorandum NMFS-NE-189.
- 594 Hart, D.R. and Chute, A.S., 2009. Estimating von Bertalanffy growth parameters from growth increment
595 data using a linear mixed-effects model, with an application to the sea scallop *Placopecten magellanicus*.
596 *ICES J Mar. Sci.* 66(10):2165-2175.
- 597 Harris, B.P., and Stokesbury, K.D.E. 2006. Shell growth of sea scallops (*Placopecten magellanicus*) in the
598 southern and northern Great South Channel, USA. *ICES J. Mar. Sci.* 63:811-821.
- 599 Hart, D.R., Jacobson, L.D., and Tang, J. 2013. To split or not to split: Assessment of Georges Bank sea
600 scallops in the presence of marine protected areas. *Fish. Res.* 144: 74–83.
- 601 Hart, D.R., and Rago, P.J. 2006. Long-term dynamics of U.S. Atlantic sea scallop *Placopecten*
602 *magellanicus* populations. *N. Am. J. Fish. Manage.* 26: 490–501.

- 603 Hart, D.R. and Sissenwine, M.P., 2009. Marine reserve effects on fishery profits: a comment on White et
604 al. (2008). *Ecology Letters*, 12(3):E9-E11.
- 605 Hilborn, R., and Ledbetter, M. 1979. Analysis of the British Columbia salmon purse-seine fleet: dynamics
606 of movement. *J. Fish. Res. Board Can.* 36: 384–391.
- 607 Holland, D.S., and Jon G. Sutinen. 2000. Location Choice in New England Trawl Fisheries: Old Habits Die
608 Hard. *Land Econ.* 76: 133–149.
- 609 Jolly, G.M., and Dickson, J.M. 1983. The problem of unequal catchability in mark–recapture estimation
610 of small mammal populations. *Can. J. Zool.* 61: 922–927.
- 611 Kerr, L.A., Cadrin, S.X., and Secor, D.H. 2010. Simulation modelling as a tool for examining the
612 consequences of spatial structure and connectivity on local and regional population dynamics. *ICES*
613 *Journal of Marine Science: Journal du Conseil* 67(8): 1631–1639.
- 614 McGilliard, C.R., Punt, A.E., Methot Jr, R.D. and Hilborn, R. 2014. Accounting for marine reserves using
615 spatial stock assessments. *Can. J. Fish. Aq. Sci.* 72(2):262-280.
- 616 McShane, P.E., and Naylor, J.R. 1995. Small-scale spatial variation in growth, size at maturity, and
617 yield-and egg-per-recruit relations in the New Zealand abalone *Haliotis iris*. *New Zeal. J. Mar. Fresh.* 29:
618 603–612.
- 619 Merrill, A., and Posgay, J. 1964. Estimating the natural mortality rate of the sea scallop (*Placopecten*
620 *magellanicus*). *International Commission for the Northwest Atlantic Fisheries Research Bulletin* 1: 88–
621 106.

- 622 Morgan, L.E., Wing, S.R., Botsford, L.W., Lundquist, C.J., and Diehl, J.M. 2000. Spatial variability in red
623 sea urchin (*Strongylocentrotus franciscanus*) recruitment in northern California. Fish. Oceanogr. 9: 83–
624 98.
- 625 Naidu, K. S., and Robert, G. 2006. Fisheries sea scallop, *Placopecten magellanicus*. In *Scallops: Biology,*
626 *Ecology and Aquaculture*, 2nd edn, pp. 869–905. Ed. by S. E. Shumway, and G. J. Parsons. Elsevier,
627 Boston. 1500 pp.
- 628 NEFSC (Northeast Fisheries Science Center). 2010. Assessment of Atlantic sea scallops. In *Report of the*
629 *50th Northeast Regional Stock Assessment Workshop (50th SAW) Stock Assessment Committee (SARC)*
630 *Consensus Summary of Assessments NEFSC Reference Document 10-17*, pp. 393-708.
- 631 NEFSC (Northeast Fisheries Science Center). 2014. Assessment of Atlantic sea scallops. In *Report of the*
632 *50th Northeast Regional Stock Assessment Workshop (50th SAW) Stock Assessment Committee (SARC)*
633 *Consensus Summary of Assessments NEFSC Reference Document 14-09*, pp. 483-631.
- 634 NMFS (National Marine Fisheries Service). 2013. In *Fisheries of the United States 2013*. Ed. by Lowther
635 A., Liddel M. Current Fishery Statistics No. 2013. 130 pp.
- 636 Orensanz, J.M., and Jamieson, G.S. 1998. The assessment and management of spatially structured
637 stocks: an overview of the North Pacific Symposium on Invertebrate Stock Assessment and
638 Management. *Canadian Special Publication of Fisheries and Aquatic Sciences*: 441–459.
- 639 Orensanz, J.M., Parma, A.M., Turk, T., and Valero, J. 2006. Chapter 14 Dynamics, assessment and
640 management of exploited natural populations. In *Scallops: Biology, Ecology and Aquaculture*. Elsevier.
641 pp. 765–868.

- 642 Palmer M. C., Wigley S. E. 2009. Using positional data from vessel monitoring systems to validate the
643 logbook-reported area fished and the stock allocation of commercial fisheries landings. N. Am. J. Fish.
644 Manage. 29: 928–942.
- 645 Paloheimo, J.E., and Dickie, L.M. 1964. Abundance and fishing success. Rapp. P.-V. Réun. Cons. Int.
646 Explor. Mer. 155: 152–163.
- 647 Quinn, T.J., and Deriso, R.B. 1999. Quantitative Fish Dynamics. Oxford University Press, New York.
- 648 Ralston, S. and O'Farrell, M.R. 2008. Spatial variation in fishing intensity and its effect on yield. Can. J.
649 Fish. Aquat. Sci. 65(4):588-599.
- 650 Sampson, D.B. and Scott, R.D., 2011. A spatial model for fishery age-selection at the population level.
651 Can. J. Fish. Aquat. Sci. 68(6):1077-1086.
- 652 Schnute, J.T., Boers, N. and Haigh, R. 2015. PBSmapping: Mapping Fisheries Data and Spatial Analysis
653 Tools. R package version 2.69.76. <http://CRAN.R-project.org/package=PBSmapping>.
- 654 Selgeby, J.H. 1982. Decline of lake herring (*Coregonus artedii*) in Lake Superior: an analysis of the
655 Wisconsin herring fishery, 1936-78. Can. J. Fish. Aquat. Sci. 39: 554–563.
- 656 Serchuk, F.M., Wood, P.W., Posgay, J.A., and Brown, B.E. 1978. Summary and review of the 1978
657 assessment and status of sea scallop (*Placopecten magellanicus*) populations off the northeast coast of
658 the United States. NEFSC Reference Document 78-45. 25 pp.
- 659 Shank, B.V., Hart, D.R., and Friedland, K.D. 2012. Post-settlement predation by sea stars and crabs on
660 the sea scallop in the Mid-Atlantic Bight. Mar. Ecol. Prog. Ser. 468: 161.

- 661 Stokesbury, K.D.E. 2002. Estimation of Sea Scallop Abundance in Closed Areas of Georges Bank, USA.
662 Transactions of the American Fisheries Society 131(6): 1081–1092.
- 663 Sullivan, P.J., Lai, H.-L., and Gallucci, V.F. 1990. A catch-at-length analysis that incorporates a stochastic
664 model of growth. Can. J. Fish. Aquat. Sci. 47: 184–198.
- 665 Truesdell, S.B., Bence, J.R., Syslo, J.S., and Ebener, M.P. Estimating multinomial effective sample size in
666 catch-at-age and catch-at-size models. Fish. Res. *In press*.
- 667 Truesdell, S.B., Hart, D.R., and Chen, Y. 2016. Effects of spatial heterogeneity in growth and fishing effort
668 on yield-per-recruit models: an application to the US Atlantic sea scallop fishery. ICES J. Mar. Sci. 73:
669 1062–1073.
- 670 Wilberg, M.J., and Bence, J.R. 2006. Performance of time-varying catchability estimators in statistical
671 catch-at-age analysis. Can. J. Fish. Aquat. Sci. 63: 2275–2285.
- 672 Wilen, J.E., Smith, M.D., Lockwood, D., and Botsford, L.W. 2002. Avoiding surprises: incorporating
673 fisherman behavior into management models. Bull. Mar. Sci. 70: 553–575.
- 674

675 Table 1. Standardized root mean square error for fully selected fishing mortality (F), catch, abundance
 676 of scallops > 100 mm (N) and number of recruits. For each scenario at selected levels of heterogeneity
 677 (ν), this calculation was aggregated all years across all iterations. The root mean square error was
 678 standardized by the mean of the simulated values.

Scenario	ν	F	Catch	N	Recruits
IDF	0	0.030	0.023	0.029	0.179
	0.5	0.046	0.040	0.045	0.178
	1	0.084	0.097	0.093	0.183
	1.5	0.118	0.147	0.125	0.186
WIDF	0	0.030	0.023	0.029	0.179
	0.5	0.047	0.041	0.046	0.178
	1	0.084	0.098	0.093	0.184
	1.5	0.115	0.144	0.123	0.187

679

Figure 1. Example of simulation scenarios. The study area is comprised of 10-minute by 10-minute squares, each of which is a unit in the simulation model. The four scenarios represent examples of combinations of available biomass (the color of the squares; log biomass shown to make the process differences more apparent) and fully selected fishing mortality (the size of the circles). The inset plots give the distribution of fully selected fishing mortality, which depends on v . The two crosses that are visible represent the ports of Montauk, New York and Cape May, New Jersey. In panel A fishing

687 mortality is homogeneous ($\nu=0$), in panel B fishing mortality follows an ideal free distribution with $\nu=1$
688 and in panel C fishing mortality follows a weighted ideal free distribution with $\nu=1$. In C, the WIDF,
689 larger fishing mortalities can be seen closer to Montauk, NY (the cross near the top of the figure).

Draft

690

691

692

693

694

695

Figure 2. Example comparison for the CAS model fully selected fishing mortality estimates at different levels of ν . The CAS model begins to overestimate fishing mortality as ν increases. This plot is an example of a single simulation-CAS model pair (i.e., one iteration) under the scenarios $\nu = 0$, $\nu = 0.5$, $\nu = 1$ and $\nu = 1.5$. Not all model runs produced overestimated fishing mortality in every year, but this example demonstrates the general average trend of overestimation (see Fig. 5).

696

697 Figure 3. Example comparison for the CAS model abundance estimates for scallops > 100 mm shell
 698 height at different levels of v . The CAS model begins to underestimate abundance as v increases. This
 699 plot is an example of a single simulation-CAS model pair (i.e., one iteration) under the scenarios $v = 0$,
 700 $v = 0.5$, $v = 1$ and $v = 1.5$. Not all model runs produced underestimated abundance in every year, but
 701 this example demonstrates the general average trend of underestimation (see Fig. 5).

Figure 4. Example comparison for the CAS model estimates of proportions-at-size in the population at $v = 0$, $v = 0.5$, $v = 1$ and $v = 1.5$. The height of the bars represents the proportion of individuals in a size class from the simulation minus the estimated proportion from the assessment. Shown here is year 25 of a 40-year simulation (which is representative of the general pattern). As v increases there is an accumulation of scallops in the plus group that is not well estimated by the CAS model.

708

709 Figure 5. Bias in the catch-at-size model estimates for fishing mortality and abundance of scallops 100
 710 mm and over. The vertical lines represent the 25th – 75th percentiles of the average CAS model
 711 deviations from the simulation (i.e., the average of the CAS estimate – the “true” value from simulation)
 712 for each iteration in every year. As v increases, the CAS model tends to overestimate fishing mortality
 713 and underestimate abundance for the IDF and WIDF scenarios. Each scenario (e.g., IDF and $v = 0.5$)
 714 was comprised of 50 iterations of the simulated population and stock assessment model.

Figure 6. Bias in the catch-at-size model estimates for fishing mortality and abundance of scallops 100 mm and over under the alternative scenario A1 where the likelihood weighting for the commercial catch data was reduced. The vertical lines represent the 25th – 75th percentiles of the average CAS model deviations from the simulation (i.e., the average of the CAS estimate – the “true” value from simulation) for each iteration in every year. As v increases, the CAS model tends to overestimate fishing mortality

721 and underestimate abundance for the IDF and WIDF scenarios. Each scenario (e.g., IDF and $\nu = 0.5$)
722 was comprised of 50 iterations of the simulated population and stock assessment model.

723

Draft

Figure 7. Bias in the catch-at-size model estimates for fishing mortality and abundance of scallops 100 mm and over under the alternative scenario A2 where the likelihood weighting for the commercial catch composition data (i.e., effective sample size) was reduced. The vertical lines represent the 25th – 75th percentiles of the average CAS model deviations from the simulation (i.e., the average of the CAS estimate – the “true” value from simulation) for each iteration in every year. As v increases, the CAS model tends to slightly underestimate both fishing mortality and abundance in both the IDF and WIDF

731 scenarios. Each scenario (e.g., IDF and $\nu = 0.5$) was comprised of 50 iterations of the simulated
732 population and stock assessment model.

733

Draft

Figure 8. Bias in the Beverton-Holt estimator of fishing mortality under homogeneous fishing mortality (top) and heterogeneous fishing mortality with $v=1.5$ and under the IDF scenario (bottom). The solid line represents the true average fishing mortality.