

Testing the performance of a spatially explicit tag-integrated stock assessment model of yellowtail flounder (*Limanda ferruginea*) through simulation analysis

Daniel R. Goethel, Christopher M. Legault, and Steven X. Cadrian

Abstract: In any stock assessment application, the implicit assumptions regarding spatial population structure must be carefully evaluated. Tag-integrated models offer a promising approach for incorporating spatial structure and movement patterns in stock assessments, but the complexity of the framework makes implementation challenging and the appraisal of performance difficult. A flounder-like fishery was simulated to emulate the metapopulation dynamics of the three yellowtail flounder (*Limanda ferruginea*) stocks off New England, and the robustness of spatially explicit tag-integrated models were compared with closed population assessments. Different movement parametrizations and data uncertainty scenarios were simulated, while the ability of the tag-integrated model to estimate reporting rate and time-varying movement were also evaluated. Results indicated that the tag-integrated framework was robust for the simulated fishery across a wide range of connectivity levels and that tag reporting rates were accurately estimated. Closed population models also demonstrated limited error. Therefore, spatially explicit approaches may not always be warranted even when regional connectivity is occurring, but tag-integrated models can provide improved parameter estimates when reliable tagging data are available. Tag-integrated models also serve as valuable tools for informing spatially explicit operating models, which can then be used to evaluate the assumptions and performance of closed population models.

Résumé : Dans toute application d'évaluation des stocks, les hypothèses implicites concernant la structure spatiale de la population doivent faire l'objet d'une évaluation soigneuse. Si les modèles intégrant les données de marquage constituent une approche prometteuse pour incorporer la structure spatiale et les motifs de déplacement aux évaluations des stocks, la complexité de leur cadre en rend l'application compliquée et l'évaluation de la performance difficile. Une pêche s'apparentant à la pêche à la limande a été simulée pour reproduire la dynamique de métapopulation des trois stocks de limandes à queue jaune (*Limanda ferruginea*) au large de la Nouvelle-Angleterre, et la robustesse de modèles spatialement explicites intégrant le marquage a été comparée avec des évaluations de population fermée. Différentes paramétrisations des déplacements et différents scénarios d'incertitude des données ont été simulés, et la capacité du modèle intégrant le marquage à estimer les taux de déclaration et les déplacements variables dans le temps ont aussi été évalués. Les résultats indiquent que le cadre intégrant les données de marquage est robuste pour la pêche simulée pour une grande fourchette de niveaux de connectivité et que les taux de déclaration des marques étaient estimés avec exactitude. Les modèles de population fermée présentaient également des erreurs restreintes. Ainsi, si les approches spatialement explicites ne sont pas toujours nécessaires même quand il y a connectivité régionale, les modèles intégrant le marquage peuvent toutefois fournir de meilleures estimations des paramètres quand des données de marquage fiables sont disponibles. Les modèles intégrant les données de marquage constituent également de bons outils pour alimenter les modèles d'exploitation spatialement explicites, qui peuvent ensuite être utilisés pour évaluer les hypothèses et la performance des modèles de population fermée. [Traduit par la Rédaction]

Introduction

Spatial population structure and movement patterns play a critical role in the regulation of abundance (Cadrin and Secor 2009). Attempting to incorporate the intricacies of fine-scale population structure into spatially explicit stock assessment models has proven difficult because of the increased data demands (Cope and Punt 2011; Berger et al. 2012). “Spatially referenced” assessments that implicitly account for population structure by allowing for spatially varying selectivity (i.e., the fleets-as-areas approach) reduce data demands, but are only a proxy for the true spatial processes (Berger et al. 2012; Hurtado-Ferro et al. 2014). Tag-integrated models, which include tagging data in the objective function, provide one approach for directly incorporating spatial structure

and movement (Maunder 1998; Goethel et al. 2011). However, the tag-integrated framework is relatively new and has only been applied to a few case studies (e.g., MULTIFAN-CL, Hampton and Fournier 2001; Integrated Tagging and Catch-at-Age Analysis (ITCAAN), Maunder 1998; Multi-stock Age-Structured Tag-integrated model (MAST), Taylor et al. 2011; Stock Synthesis, Methot and Wetzel 2013; see Goethel et al. 2011 for a review of tag-integrated approaches).

Stock assessment models require extensive validation to determine reliability and performance. Model evaluation is commonly achieved by developing an “operating model” to simulate data and applying the stock assessment on this “observed” data (Kerr and Goethel 2014). The operating model emulates the real-world system of interest by simulating the system forward from an ini-

Received 21 May 2014. Accepted 9 December 2014.

Paper handled by Associate Editor Michael Wilberg.

D.R. Goethel and S.X. Cadrin. University of Massachusetts, School for Marine Science and Technology, 200 Mill Road, Fairhaven, MA 02719, USA.
C.M. Legault. National Marine Fisheries Service, Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543, USA.

Corresponding author: Daniel R. Goethel (e-mail: Daniel.Goethel@noaa.gov).

tial state using stochastic processes and observation error. Estimates of population parameters from application of the stock assessment model to the multiple simulated data sets can then be summarized and compared with the “true” values used by the operating model (Hilborn and Walters 1992). Attempting to evaluate all sources of uncertainty in a given model is impossible, but with the development of any stock assessment method it is important to determine the range of situations where it can be satisfactorily applied (Kolody et al. 2004).

Spatial population structure and movement patterns are often drivers of population dynamics, in conjunction with the influences of fishing and the environment, but vary widely among species (Hilborn and Walters 1992). Marine spatial structure can take many forms (e.g., natal homing or metapopulations) and impacts how components of a population interact in space and time (e.g., levels of interbreeding among units and spatial overlap throughout the year; Goethel et al. 2011). Most assessment models assume that the population unit is closed to immigration and emigration while implying homogeneity within the unit (Guan et al. 2013). Despite the importance of spatial structure in determining the underlying dynamics of a species, relatively little focus has been given to the impact of misspecifying population structure in stock assessment models (Smedbol and Stephenson 2001; Goethel et al. 2011).

Considering the importance of regional connectivity, evaluating the ability of assessment models to explicitly estimate movement is critical (Goethel et al. 2011). Early simulations with tag-integrated models tested the feasibility of estimating movement and compared error with closed population models that did not include mixing (e.g., Maunder 1998, 2001; Porch et al. 1998). Hulson et al. (2011, 2013) simulation-tested a tag-integrated model by exploring error from different tagging protocols and assumed movement dynamics. Although these studies demonstrated the feasibility of applying spatially explicit tag-integrated assessments, model robustness and critical sources of error have yet to be thoroughly investigated (Maunder 1998; Porch et al. 1998; Hulson et al. 2011, 2013). For instance, Hulson et al. (2011, 2013) only investigated the performance of tag-integrated models for a single, spatially structured population, while the feasibility of estimating reporting rates was not thoroughly explored by any of these simulation studies.

Our goal was to explore how error is manifest when an assessment model ignores metapopulation structure and whether error can be reduced by allowing the model to estimate movement or by including tagging information. Simulation analysis was used to emulate metapopulation dynamics of yellowtail flounder in the New England region. Multiple estimation models were run on the simulated data, and each model differed in either population structure assumptions or data sources that were utilized (e.g., tagging data). A total of five simulation scenarios (i.e., groups of analyses) were investigated. The primary focus was to better understand the performance of the spatially explicit tag-integrated model developed by Goethel et al. (2015), which incorporated all available data sources for yellowtail flounder off New England and explicitly estimated movement and reporting rates.

Methods

Operating model

The operating model was spatially explicit and simulated the dynamics of the three yellowtail flounder stocks off the coast of New England (i.e., Cape Cod – Gulf of Maine, Georges Bank, and Southern New England – Mid-Atlantic; Fig. 1). Stocks were defined by Cadrin (2010) and were used as the basis for current stock assessments (i.e., independent closed population models for each

stock unit; NEFSC 2012a, 2012b; Legault et al. 2014). For simplicity we retained the “stock” notation to refer to either an isolated spawning component in closed population models (i.e., a model that assumes no movement among units) or a subpopulation within a metapopulation (i.e., for spatially explicit models that allowed movement; see Kritzer and Sale 2004 for a discussion of marine metapopulations). In general, our definition of “stock” is that of a spatially delineated population unit within which all individuals have identical demographics.

Best-fit parameter estimates from Goethel et al. (2015) were used to develop the operating model, but movement and reporting rates were input based on exploratory model runs and did not represent the best-fit values (see below). Time series of recruitment and fishing mortality were used to simulate data, because yellowtail flounder dynamics demonstrated periodic extreme recruitment events (e.g., the 1987 year-class in the Southern New England – Mid-Atlantic stock), which would be difficult to emulate with process error alone (Deroba et al. 2015). Life history data were taken from the most recent yellowtail flounder stock assessments (NEFSC 2012a, 2012b; Legault et al. 2014). Important model settings for each of the simulation scenarios are provided in Table 1.

Metapopulation dynamics were simulated through time-invariant connectivity rates among stocks. Movement occurred instantaneously at the end of the year and represented the proportion of the stock that moved. Connectivity rates were approximately 0.2 from each stock area to each of the other two stock areas, representing a yearly residency rate around 60%. Although these values were greater than movement estimates for yellowtail flounder (i.e., values were below 10% for the tag-integrated application of Goethel et al. 2015), estimates solely from tagging data suggested that yearly movement rates around 20% were not unreasonable (Cadrin 2010; Wood and Cadrin 2013). In addition, levels of residency around 60% struck a balance between the minimal impact on closed population model performance seen at residency rates of 80% (see the online Supplementary Material, Figs. S1–S9 for results with lower connectivity rates¹) and the extreme dynamics of highly migratory species (Taylor et al. 2011).

Stock (j) abundance (N) (by year (y) and age (a)) starting from age-2 (age-1 recruitment was input to the model) to the plus group (A) at age-6 was simulated based on yellowtail flounder life history information (i.e., maturity (m), mass (w), and a fixed natural mortality of $M = 0.2$) and the best-fit parameter estimates (i.e., yearly fishing mortality (F), selectivity (v), and recruitment (R)) along with input values of box-transfer movement (T) between stocks:

$$(1) \quad N_{j,y,a} = \sum_k T_{k,j} N_{k,y-1,a-1} e^{[-(v_{k,y-1,a-1} F_{k,y-1} + M)]}, \quad 1 < a < A$$

$$N_{j,y,A} = \sum_k T_{k,j} N_{k,y-1,A-1} e^{[-(v_{k,y-1,A-1} F_{k,y-1} + M)]} + \sum_k T_{k,j} N_{k,y-1,A} e^{[-(v_{k,y-1,A} F_{k,y-1} + M)]}, \quad a = A$$

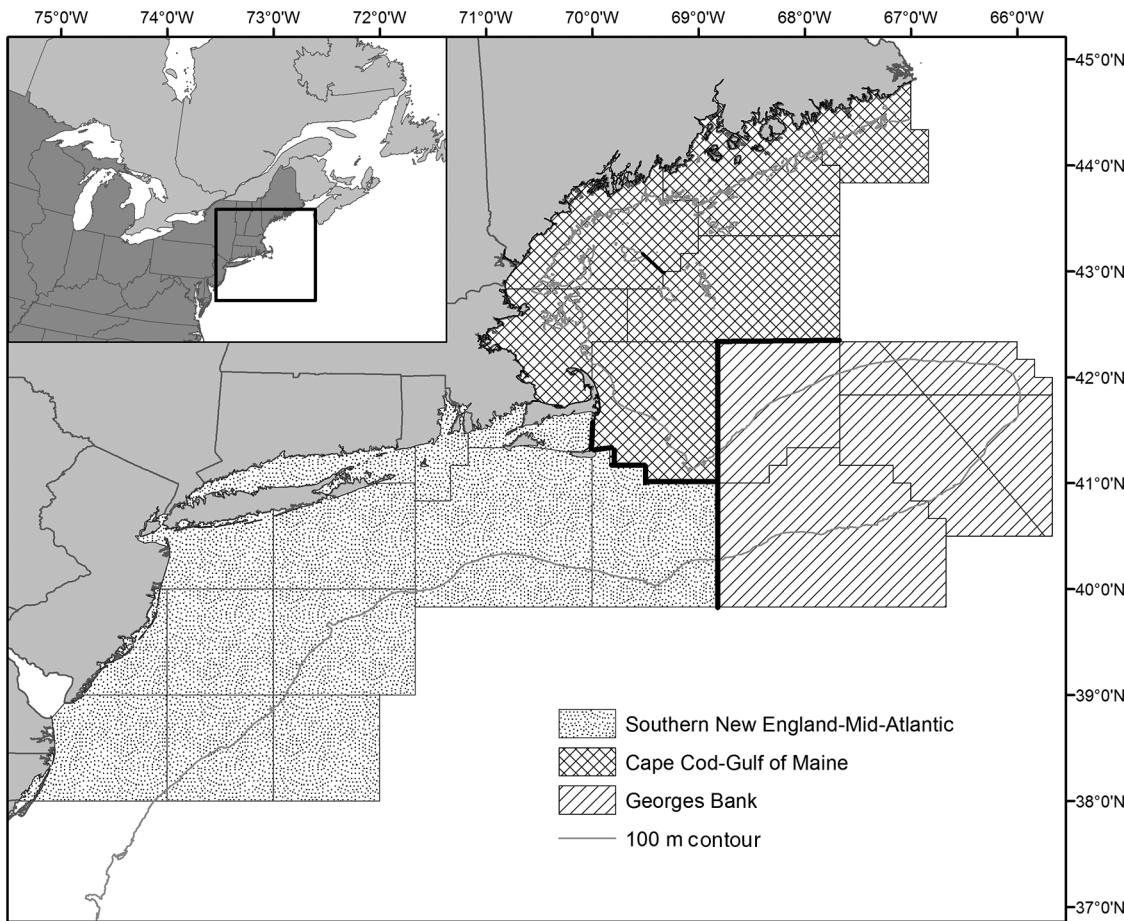
Similarly, spawning-stock biomass (SSB) was calculated at the time of spawning (t_{spawn}) by:

$$(2) \quad SSB_{j,y} = \sum_{a=1}^A w_{j,y,a} m_{j,y,a} N_{j,y,a} e^{[-(v_{j,y,a} F_{j,y} + M) \times t_{\text{spawn}}]}$$

An alternate form of movement was also simulated in Scenario 5. For this scenario, movement was assumed to be

¹Supplementary data are available with the article through the journal Web site at <http://nrcresearchpress.com/doi/suppl/10.1139/cjfas-2014-0244>.

Fig. 1. Model domain for the simulation including the boundaries for the three yellowtail flounder stocks off New England (adapted from Goethel et al. 2015).



density-dependent and was no longer constant over time. The simulated yearly total emigration rate out of a stock area was a logistic function that increased with SSB in the stock area in the previous year:

$$(3) \quad T_{j,k,y} = \Omega_{\text{Rel},k,i,y} \frac{T'_j}{1 + Ae^{-\frac{\ln(A)}{B_j^*} \times SSB_{j,y-1}}}$$

A was a logistic model parameter and B^* was the x -axis point of inflection in the logistic curve. T' was the maximum total emigration rate (saturation level in T) out of stock j , and emigrating fish were partitioned between each of the other stocks (i.e., units k and i) based on the relative biomass (Ω_{Rel}) between those stocks:

$$(4) \quad \Omega_{\text{Rel},k,i,y} = 1 - \left(\frac{\frac{SSB_{k,y-1}}{B_k^*}}{\frac{SSB_{k,y-1}}{B_k^*} + \frac{SSB_{i,y-1}}{B_i^*}} \right)$$

The relative biomass apportioned emigration by comparing the density ratios (SSB/B^*) of the destination stocks. Ω_{Rel} was formulated in a way that forced more fish to move to the less densely inhabited stock area.

The B^* parameter was the critical density-dependent term that acted as a carrying-capacity or habitat suitability proxy. The ratio of SSB to B^* in a given year determined both emigration and

immigration levels. The smaller the ratio, the more suitable the stock area would appear to fish immigrating from other stock units and the less likely a fish would be to emigrate. The formulation was similar to MacCall's (1990) basin theory, which suggested that stock growth became limited as density increased. Therefore, when a stock became densely inhabited, fish would move out of a stock in the search for a less dense area. Stock-specific B^* values allowed each stock to have different threshold densities, which implied that each stock had a unique, inherent carrying capacity.

The parameter values for the density-dependent movement model are provided in Table 2, while Fig. 2 illustrates the density-dependent movement curves for each stock and the actual movement rates used for the simulation. From the application of the tag-integrated model to yellowtail flounder (Goethel et al. 2015), it was apparent that regional biomass was dominated by the Georges Bank stock, with the Cape Cod – Gulf of Maine stock being only a fraction of the size of the others. For this reason, the Cape Cod – Gulf of Maine unit was assumed to have a much smaller B^* . Maximum emigration values also varied by stock, with the Georges Bank and Cape Cod – Gulf of Maine stocks assumed to have similar levels. The Southern New England – Mid-Atlantic stock had a higher maximum emigration rate based on the assumption that the observed extreme year-class in this stock would lead to large emigration events. However, the entire formulation for density-dependent movement was hypothetical, and all parameter values were approximations not based on real world data.

Table 1. Attributes of the simulation scenarios and estimation models.

Simulation scenario	Simulated movement	Estimation model	Assumed movement (no. of estimated parameters)	Assumed reporting rate	Years of tagging data	Data uncertainty	
Base model (Scenario 1)	Constant, $T \sim 0.2$	CLOSED	T = 0 (0) Constant T (6)	NA	NA	Base variance	
		TAG NM		EST	11		
		SPAT CNST		NA	NA		
		TAG CNST		EST	11		
Tagging methodology (Scenario 2)	Constant, $T \sim 0.2$	TAG CNST	Constant T (6)	EST	4	Base variance	
Reporting rate treatment (Scenario 3)	Constant, $T \sim 0.2$	TAG CNST $\beta = 0.5$	Constant T (6)	Fixed at 0.5	11	Base variance	
Data uncertainty (Scenario 4)	Constant, $T \sim 0.2$	CLOSED	T = 0 (0)	NA	NA	Medium variance	
		SPAT CNST		EST	11		
		TAG CNST	Constant T (6)	NA	NA	High variance	
		CLOSED		NA	NA		
		SPAT CNST	Constant T (6)	EST	11	High tag variance	
		TAG CNST		EST	11		
Movement parametrization (Scenario 5)	Constant, $T \sim 0.2$	TAG 11TB	11 T time blocks (66)	EST	11	Base variance	
		TAG RW	Estimate random walk in T (126)				
		CLOSED	T = 0 (0)	NA	NA		
		TAG NM		EST	11		
		SPAT CNST	Constant T (6)	NA	NA	Base variance	
		SPAT 11TB	11 T time blocks (66)				
		SPAT RW	Estimate random walk in T (126)	EST	11		
		SPAT YRLY	Estimate yearly T (126)				
		TAG CNST	Constant T (6)	EST	11		
		TAG 11TB	11 T time blocks (66)				
		TAG RW	Estimate random walk in T (126)	EST	11		
		TAG YRLY	Estimate yearly T (126)				

Note: The estimation model codes used throughout the manuscript are as follows: CLOSED, closed population model; TAG NM, tag-integrated model with no movement; SPAT CNST, spatial model without tagging data and constant movement; TAG CNST, spatially explicit tag-integrated model with constant movement; TAG CNST $\beta = 0.5$, spatially explicit tag-integrated model with constant movement and reporting rate fixed at 0.5. The movement model codes for Scenario 5 are as follows: 11TB, 11 movement time blocks (i.e., biennial movement); RW, a random walk in movement; YRLY, freely estimated yearly movement. Other notations are as follows: T, movement rate; β , reporting rate; NA, not applicable; EST, parameter estimated; ESS, tagging effective sample size.

Table 2. Parameter values for simulated density-dependent movement given by eqs. 3 and 4.

Parameter	Stock	Value
T'	CC	0.30
	GB	0.33
	SN	0.67
B*	CC	3702
	GB	8987
	SN	8587
A	All	9

Note: T' represents the maximum total emigration rate, B^* is a carrying capacity proxy (see text for a detailed explanation), and A is a logistic function parameter that is uniform across stocks. B^* is in metric tonnes and T' is a proportion. CC, Cape Cod – Gulf of Maine; GB, Georges Bank; SN, Southern New England–Mid-Atlantic.

Data generation

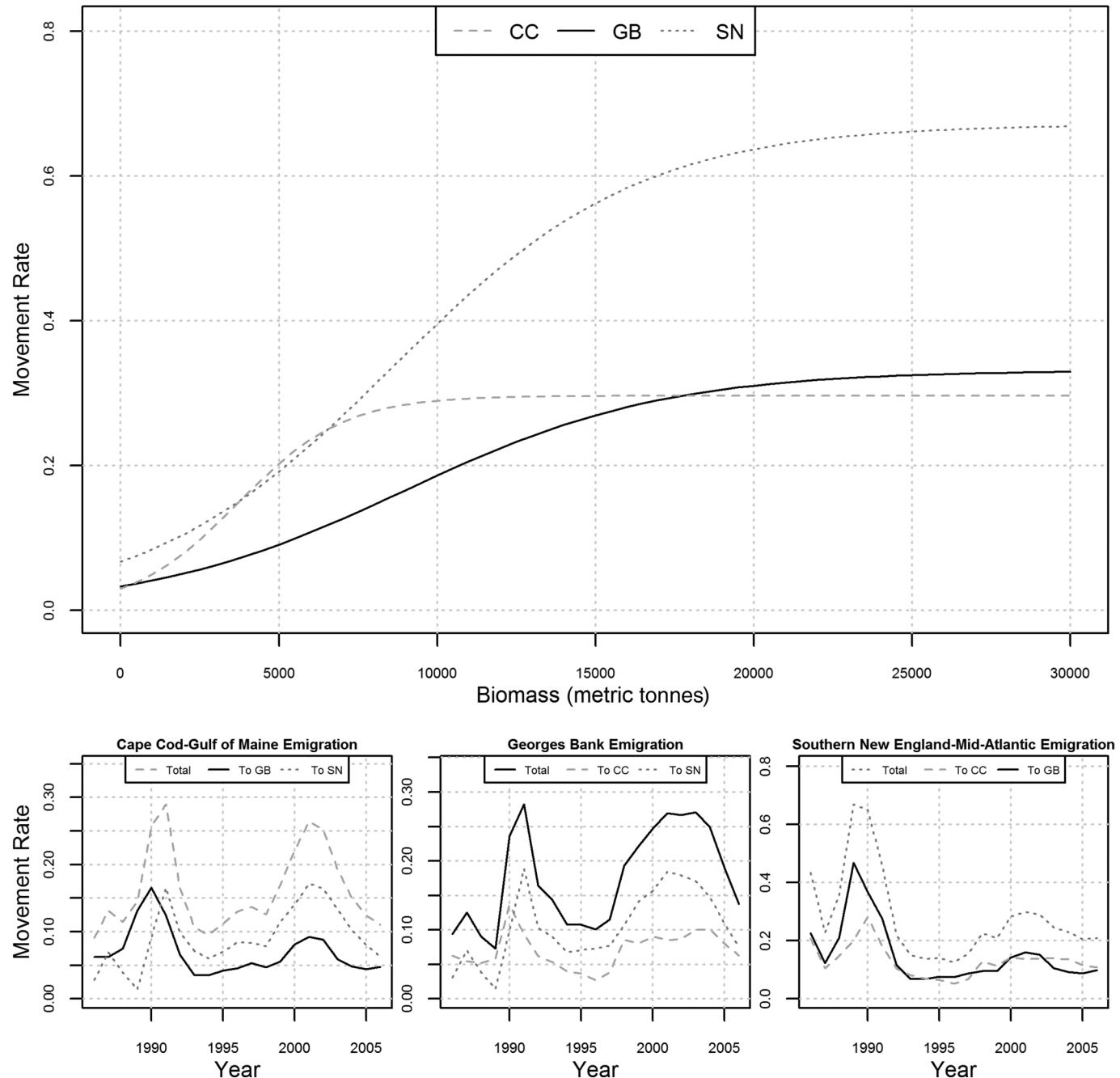
Stochastic processes in the operating model occurred in the form of observation error in the simulated data. Simulated data sources included region-specific (i) age compositions from the catch; (ii) total yield; (iii) spring and fall fishery-independent survey biomass; (iv) spring and fall fishery-independent survey age compositions; and (v) a region-wide mark–recapture tagging study.

Simulating tagging data involved modeling tag abundance (n) by cohort (l) (i.e., combination of release area and year) through time and calculating tag recaptures (r) based on the model parameters and a value for tag reporting rate (β):

$$(5) \quad n_{j,y}^l = \sum_{k=j}^i T_{k,j} n_{k,y-1}^l e^{-(F_{k,y-1} + M)} \\ r_{j,y}^l = n_{j,y}^l \frac{F_{j,y} \{1 - e^{-(F_{j,y} + M)}\}}{F_{j,y} + M}$$

Both the simulation and estimation models assumed that all tagged fish were fully selected, because the tagging study that the simulation emulated utilized commercial fishing boats and gear with steep logistic selectivity curves (i.e., near 100% selectivity by age-3; see Goethel et al. 2015). Also, no age information was available from tagging data to compare selectivity assumptions between tagged and untagged fish. Because the assumptions are congruent between the estimation and simulation models, no model misspecification occurs owing to the way in which selectivity is modeled for tagged fish. For the tagging submodel it was also assumed that released fish were immediately well-mixed with the untagged population, tag-induced mortality was negligible, and there was no tag loss. Simulated tagging protocols and assumptions were implemented to mimic those of the real world

Fig. 2. Simulated density-dependent movement curves used in Scenario 5. Rates represent the fraction of the population that moves. The top plot illustrates the theoretical curves for a range of biomass. The bottom plots show the emigration rates used in the simulation based on true biomass levels and relative biomass ratios. Note that plots have different scales. CC, Cape Cod – Gulf of Maine; GB, Georges Bank; SN, Southern New England – Mid-Atlantic.



yellowtail flounder tagging data, which were analyzed by Cadrin (2006), Alade (2008), and Wood and Cadrin (2013).

The actual tagging study for yellowtail flounder carried out by Cadrin (2006) was emulated by releasing 10 000 tags per year proportional to the relative fall survey biomass in each stock area. The number of releases was chosen based on the mean number of tags released in the yellowtail flounder tagging program (Cadrin 2006). Each stock area had an independent time-invariant reporting rate with a value that ranged from 0.2 to 0.3. Similar to simulated movement rates, the values were slightly above those estimated from the tag-integrated model (Goethel et al. 2015), but balanced the relatively high estimate of 0.55 from the high reward

tag ratio method of Cadrin (2006). Tagging was implemented for half of the time-series (i.e., 11 years) beginning at the midpoint (i.e., 1996), and recoveries were possible over the same period.

Each of the data sources were simulated based on the modeling framework outlined in Goethel et al. (2015). Observation error was then incorporated into each data set based on an underlying error assumption. For yield and survey biomass, lognormal variance was assumed. For a given data source (D), error was incorporated by multiplying the true simulated data value ($\bar{x}_{D,y}$) by the lognormal random error ($e_{\text{Norm},D,y}^R$), to obtain the data ($x_{D,y}^R$) for each model run (R) and included a lognormal bias correction factor:

Table 3. Uncertainty associated with each data set used to simulate observation error in the simulation model and input as data weights in the estimation models.

Data component	Type	Base settings	Medium variance	High variance	High tag variance
Fishery age composition	ESS	100	50	50	100
Fishery yield	σ	0.05	0.1	0.2	0.05
Spring survey age composition	ESS	75	25	25	75
Spring survey biomass	σ	0.2	0.4	0.8	0.2
Fall survey age composition	ESS	75	25	25	75
Fall survey biomass	σ	0.2	0.4	0.8	0.2
Tag data	ESS	1000	1000	1000	100
Recruitment variance	σ	0.6	0.6	0.6	0.6

Note: The base settings are used in Scenarios 1–3 and Scenario 5, and the medium variance, high variance, and high tag variance values are for Scenario 4 runs. Effective sample size (ESS) and variance (σ) are constant across years and stocks. However, tagging ESS is also constant across cohorts (i.e., each tagged cohort has the same ESS given by the number in the table). The closed population model and spatial model without tagging data use the same weighting, but have no tagging component. The recruitment variance (σ) only applies to the estimation models, because recruitment streams are input into the simulation model and not simulated assuming random error.

$$(6) \quad x_{D,y}^R = \bar{x}_{D,y} e^{e_{\text{Norm},D,y}^R - 0.5\sigma_D^2}$$

$$\varepsilon_{\text{Norm},D,y}^R \sim N(0, \sigma_D^2)$$

The uncertainty in yield and abundance index data sources was controlled by the variance (σ_D^2). For catch and survey age compositions and tag proportions, multinomial error was assumed in which variance was controlled by the effective sample size (i.e., a higher effective sample size resulted in lower associated variance). In each state (b), the true proportions (or probabilities) ($\bar{p}_{D,y,b}$) and effective sample size (ESS_D) determined the multinomial distribution of the random variable ($e_{\text{Mult},D,y,b}^R$):

$$(7) \quad e_{\text{Mult},D,y,b}^R \sim \text{Mult}(\text{ESS}_D, \bar{p}_{D,y,b})$$

The perceived proportion ($\bar{p}_{D,y,b}$) in a given state (b) was the count of random variables in that state divided by the effective sample size. For age composition data, the state represented an age-class in a given year, while for tagging data a state was the combination of recapture year and location (stock area) for a given release cohort where the final state subsumed all possible states where a tag was not “seen” by scientists (i.e., the tag was not reported by a fisherman, the tagged fish died or was still at large during the terminal year of the model).

Data uncertainty resulting from observation error was determined by the variance components associated with each data set. Moderate levels of error were incorporated (Table 3), which reflected expert judgment regarding real world uncertainty due to data collection methodology. Data set variance was constant across years and stocks. For each scenario, 100 sets of data were simulated with each run using a different random number generator to encapsulate variation.

Estimation models

Four estimation models were run to assess the impact of assumed population structure and the addition of tagging information. Each estimation model was a statistical catch-at-age assessment as outlined by Goethel (2014) and Goethel et al. (2015) and was implemented in AD Model Builder (Fournier et al. 2012). Table 1 outlines the details of each estimation model, and the model abbreviations used therein are provided following the full name in the remainder of this section. A closed population model (CLOSED) was applied to demonstrate how assessing each stock individually and without connectivity with other units impacted model outputs. The tag-integrated model with no movement (TAG NM) was implemented to illustrate how tagging information improved estimates (particularly for fishing mortality parameters) within a closed population model. Both closed population

models had implicit model misspecification error caused by the incorrect population structure assumptions (i.e., both ignore movement). The TAG NM model had additional model misspecification error, because it was unable to fit to tag recaptures outside the area of release. The simulated tag recaptures from fish that moved across stock boundaries remained in the objective function, but the model was unable to match these observations because fish were not allowed to leave their natal stock (a robust likelihood method was utilized, which avoided issues with logarithms of zero values; Fournier et al. 1990). A spatially explicit model that did not include tagging data but estimated movement (SPAT CNST) was also applied to allow appraisal of the influence of tagging data in models that estimated movement. The spatially explicit tag-integrated model (TAG CNST) incorporated all available data sources while explicitly estimating connectivity levels and reporting rates.

In each of the estimation models, recruitment was calculated based on yearly estimated deviations (ε_R) from a stock-specific mean recruitment value (R_{MEAN}) and included a lognormal bias correction factor:

$$(8) \quad N_{j,y,a} = R_{\text{MEAN},j} e^{e_{R,j,y}^R - 0.5\sigma_R^2}$$

The variance component (σ_R^2) was a fixed input (value given in Table 3), which limited the extent of year-to-year deviations in recruitment. Because the operating model treated recruitment as a model input and did not simulate it from an underlying stock-recruit curve, it caused each estimation model to be misspecified (and not be able to estimate recruitment perfectly).

Population dynamics for the remainder of the ages follow eq. (1) with movement parameters set to zero for closed population models or calculated based on estimated logit-transformed movement parameters (ω), which forced a natural bound on movement between 0 and 1:

$$(9) \quad T_{k,j} = \frac{e^{\omega_{k,j}}}{\sum_{k=1}^J e^{\omega_{k,j}}}, \quad \omega_{jj} = 0$$

Alternate movement parametrizations were tested in Scenario 5, including biennial movement (i.e., 11 time blocks of movement; 11TB), a random walk in movement (RW), and freely estimated yearly movement (YRLY). The random walk was implemented using the multinomial logit transform (note that movement in the first year, $y = 1$, was not estimated in any of the estimation models):

$$(10) \quad T_{k,j,y=2} = \frac{e^{\omega_{k,j,y=2}}}{\sum_{k=1}^J e^{\omega_{k,j,y=2}}}, \omega_{jj} = 0 \quad \left| \begin{array}{l} \text{if } y = 2 \\ \omega_{k,j,y} = \omega_{k,j,y-1} \cdot \varepsilon_{RW_{k,j,y}}, \varepsilon_{RW_{j,j,y}} = 0 \\ T_{k,j,y} = \frac{e^{\omega_{k,j,y}}}{\sum_{k=1}^J e^{\omega_{k,j,y}}} \end{array} \right. \quad \left| \begin{array}{l} \text{if } y > 2 \end{array} \right.$$

The random walk was incorporated using a normal deviate within the logit transform to ensure that the random walk did not lead to movement parameters greater than one. A penalty was also included, which limited the year-to-year deviations in T :

$$(11) \quad L_{RW} = \log(\sigma_{RW}) + \frac{1}{2\sigma_{RW}^2} \sum_{k=1}^J \sum_{y=3}^Y \ln\left(\frac{T_{k,j,y}}{T_{k,j,y-1}}\right)^2$$

The variance term (σ_{RW}^2) was treated as an input with a value of 0.7 based on model stability in exploratory runs and was constant across stocks.

The total likelihood function to be minimized included components for fitting to catch proportions by age, yield, survey index proportions by age, survey biomass, tag proportions by cohort, a lognormal recruitment penalty, and a movement penalty that constrained movement estimates to reasonable values and improved model stability (see Goethel 2014). The closed population model and spatial model without tagging data did not include a component for fitting to tag data. Movement penalty terms were not included in the tag-integrated model without movement or the closed population model. All estimation models assumed the same error structure as the operating model used to simulate the data. The variance terms for each likelihood component were taken directly from the simulation model (Table 3), because error misspecification was not considered.

Evaluation of model performance

The estimation models were applied to the simulated data for each run, and error, precision, and stability were assessed for the entire scenario. Error was calculated as the difference between the parameter estimate (x) from a given run and the true value for that scenario used by the operating model. Relative error (as a percent) was the error divided by the true parameter value:

$$(12) \quad \text{Relative Error} = 100 \left(\frac{\bar{x} - x_{\text{true}}}{x_{\text{true}}} \right)$$

Evaluation of model performance focused on the comparison between the “true” and estimated time series of stock-specific spawning-stock biomass, recruitment, and fishing mortality with emphasis on terminal year estimates. The estimability of movement and reporting rate parameters were also of primary interest.

Violin plots, which overlaid a boxplot on top of a kernel density plot, were used to visualize error and precision for each scenario. Estimates were deemed inaccurate when the true value was outside the distribution of the violin plot. For parameters and derived quantities that fluctuated over time, if year-to-year variation or time trends were not detectable from the distribution given by the violin plot time series, it was deemed imprecise. In other words, if the true values varied from year to year, but the distribution of estimates overlapped from one year to the next and no changes were detectable, then the estimate was imprecise.

Model-estimated uncertainty as calculated from the inversion of the Hessian matrix in AD Model Builder for each model run was

presented in the form of a coefficient of variation (CV_{Mod}). These values were distinct from the simulation precision values (also presented as a coefficient of variation; CV_{Sim}), because they represented a direct model estimate from a single model run, whereas precision represented the distribution of a parameter estimate across all model runs for an estimation model in a given scenario. When both values were similar it indicated that the model was capable of accurately estimating uncertainty levels.

Model stability, an indicator of overparametrization and robustness, was addressed by calculating the proportion of runs that an estimation model converged. Convergence was defined as the ability of the model to obtain a stable solution (i.e., a minimum value of the objective function) and occurred when the maximum gradient component from the minimization of the objective function was less than 0.001 and both the Hessian and covariance matrices were positive-definite (i.e., the Hessian matrix was invertible and the covariance matrix was nonsingular). We considered an estimation model with a frequency of convergence greater than 95% to be performing well. Low convergence frequencies could result from the inability of a model to interpret data (as indicated by poor model fit) or parameter confounding.

Simulation scenarios

Five scenarios (groups of analyses) investigated the (i) comparative performance of spatially explicit assessment models that allow movement among units and closed population models that assume no movement occurs, (ii) effect of tagging time series length on model estimates, (iii) ability of the tag-integrated model to estimate tag reporting rate, (iv) influence of data uncertainty, and (v) error resulting from nonstationarity in movement rates.

Scenario 1: base model

The base simulation model was used to compare the relative performance of the four estimation models. The purpose of this scenario was to improve understanding of tag-integrated model performance for species that demonstrate metapopulation structure, especially in regards to estimating movement and reporting rate parameters and whether tagging data had a strong influence on parameter estimation. Additionally, we wanted to determine how ignorance of connectivity would impact closed population models. Alternate runs (see Supplementary Material¹) also looked at how different exchange rates among populations impacted the various estimation model configurations.

Scenario 2: tagging methodology

Hulson et al. (2011) demonstrated that a longer time series of tagging data improved parameter uncertainty in tag-integrated models. Scenario 2 extended their analyses for stocks, which demonstrated metapopulation structure and evaluated whether extending the tagging time series led to notable improvements in parameter estimates. The base tag-integrated model that had an 11-year time series of tagging data was compared with a tag-integrated model informed with only 4 years of tagging data spanning the final 4 years of the time series.

Scenario 3: reporting rate parametrization

Reporting rate estimates are inversely correlated with mortality (Hoenig et al. 1998), and independent estimates are often biased (e.g., when the high reward ratio method is used; Pollock et al. 2002). Scenario 3 investigated the parametrization of reporting rate by comparing the base tag-integrated model that estimated reporting rate with a model in which reporting rates were fixed around 0.5 (TAG CNST $\beta = 0.5$; true values still ranged from 0.2 to 0.3). The model with reporting rate fixed demonstrated how the tag-integrated model would respond when reporting rate was incorrectly input from an external source. Fixing reporting rate at 0.5 provides a plausible example of how the tag-integrated model might have performed if the high reward tag ratio method

Table 4. Frequency of convergence for presented estimation–simulation scenarios.

Simulation scenario	Simulation setting	Estimation model	Frequency of convergence (%)
Base model (Scenario 1)	Base model	CLOSED	96
		TAG NM	100
		SPAT CNST	95
		TAG CNST	100
Tagging methodology (Scenario 2)	4 years of tagging data	TAG CNST	100
Reporting rate treatment (Scenario 3)	Base model	TAG CNST $\beta = 0.5$	97
Data uncertainty (Scenario 4)	Medium variance	CLOSED	60
		SPAT CNST	76
		TAG CNST	77
	High variance	CLOSED	66
		SPAT CNST	61
		TAG CNST	80
	High tag variance	TAG CNST	100
		TAG 11TB	89
		TAG RW	37
Movement parametrization (Scenario 5)	Density-dependent movement	CLOSED	96
		TAG NM	100
		SPAT CNST	34
		SPAT 11TB	99
		SPAT RW	23
		SPAT YRLY	74
		TAG CNST	98
		TAG 11TB	92
		TAG RW	32
		TAG YRLY	65

Note: Convergence is defined as an estimation model being able to invert the Hessian matrix and having a nonsingular covariance matrix (i.e., both the Hessian and covariance matrices are positive-definite). The maximum gradient of the objective function must also be less than 0.001 (all models demonstrated gradients below this threshold value). Frequency of convergence is calculated as the number of runs (out of the 100 runs for each simulation scenario) that converge. Codes for the estimation models are given in Table 1.

estimate from Cadrin (2006) had been treated as a model input, while true reporting rate values were much lower.

Scenario 4: data uncertainty

An important aspect of any estimation model is how robust it is to data uncertainty. Including tagging data often improves the performance of spatially explicit models, but the degree of improvement is dependent on the uncertainty in tagging data (Maunder 1998; Hulson et al. 2011). In Scenario 4, observation error in simulated data sets was increased, and the performance of each of the base estimation models was assessed. The simulation model was run with (i) medium variance where observation error was moderately increased in all data sources except tagging; (ii) high variance where observation error was increased further for survey abundance indices and fishery yield; and (iii) high tag variance where tagging ESS was reduced to 100 with all other variances held at base levels. Observation error levels for the three data uncertainty runs are provided in Table 3.

Scenario 5: movement parametrization

Movement is unlikely to be constant. Hulson et al. (2013) illustrated that estimating time-varying movement parameters was feasible with tag-integrated models when the functional form of movement was correctly specified or a scaled random walk was assumed. However, the simulated changes in movement were unidirectional (based on climate change) or random, and the impact of density-dependent movement was not explored. Scenario 5 compared the performance of each of the base estimation models when simulated movement was density-dependent, while also testing the ability of four different movement parametrizations in the spatially explicit models (i.e., constant movement, biennial movement, a random walk, and freely estimated yearly movement). The purpose of this scenario was to investigate the error introduced when the estimation model was ignorant of non-

stationarity in movement rates, which caused the estimation models that incorporated movement to have similar process error as the closed population models. It also explored the ability of tag-integrated models to estimate time-varying movement rates.

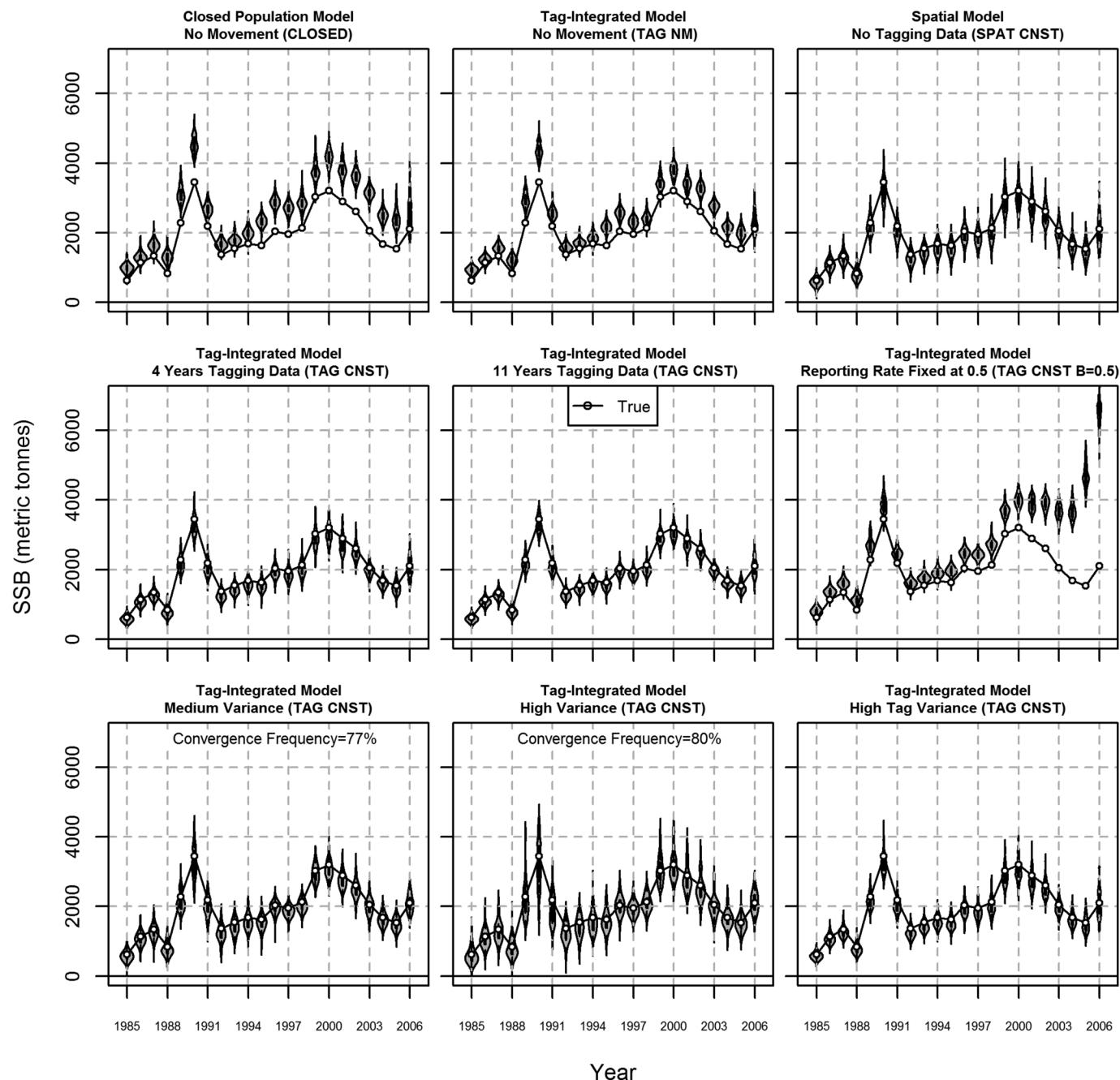
Results

Alternate violin plots for time-series estimates of fishing mortality, spawning-stock biomass, and recruitment for all scenarios and stock areas that are not presented in the main text can be found in the Supplementary Material¹.

Assessment models in Scenarios 1–3 demonstrated high stability and a general lack of residual patterns and parameter correlation (the latter two conclusions are based on visual inspections of diagnostic plots; Table 4). However, Scenarios 4 and 5 showed some important model stability issues (Table 4). Investigation of model residuals for Scenario 4 runs demonstrated that the increases in observation error caused some moderate residual patterning in survey biomass data sets. Correlations also increased, particularly among fishery and survey selectivity parameters and between mean recruitment and recruitment deviations. For the scenario with tagging effective sample size decreased to 100, correlations began to emerge among movement parameters. The convergence problems in Scenario 5 were associated with over-parametrization of the movement models. For both the random walk and yearly movement models, all of the movement parameters were highly correlated. The 11 time-block movement model had less severe correlations, but in years prior to the inclusion of tagging data strong correlations were frequent among movement parameters.

The base model runs indicated that the spatially explicit tag-integrated model (TAG CNST) was the most precise estimation model in Scenario 1 (coefficients of variation (CV_{Sim}) in terminal year parameter estimates, excluding recruitment, around 10%)

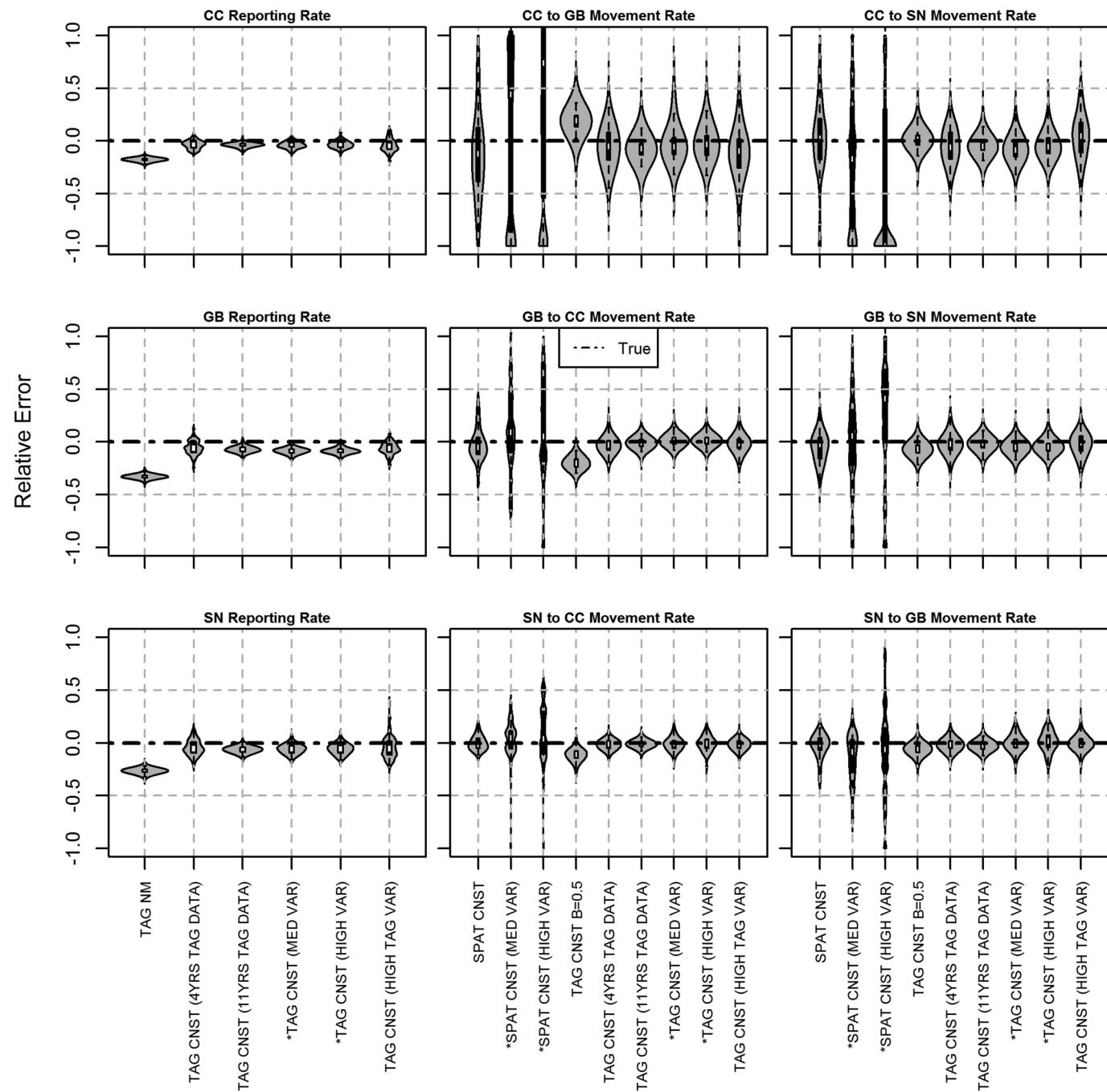
Fig. 3. Time-series estimates of Cape Cod – Gulf of Maine spawning-stock biomass for Scenarios 1–4 with constant simulated movement around 0.2. True values are given by the black line with white dots. All models had a frequency of convergence greater than 95% unless otherwise noted on the plot. Model codes are provided in Table 1.



and demonstrated the most accurate estimates of many parameters, including reporting rate and movement (mean percent error in terminal year parameter estimates, except recruitment, around 5%; Figs. 3–5; Tables 5–6). The closed population model (CLOSED) demonstrated larger error than the spatially explicit tag-integrated models in most parameter estimates especially for the Cape Cod – Gulf of Maine stock (mean percent error of +23% in terminal year spawning-stock biomass). The tag-integrated model without movement (TAG NM) provided similar results as the closed population model, but precision was higher (CV_{Sim} around 15% versus 20% for the closed population model) with less error in spawning-stock biomass estimates (mean percent error around 10%). However, terminal year fishing mortality estimates demon-

strated higher error (mean percent error of 25%–35%), which was caused by an underestimate of tag reporting rate in all stocks. The spatial model without tagging data (SPAT CNST) had lower error than the closed population models (mean percent error around 5%–10%, excluding recruitment), but was more imprecise than the tag-integrated model (CV_{Sim} around 20%, excluding recruitment, with movement CV_{Sim} reaching 60%). Terminal year recruitment estimates were the most difficult parameters to assess, with all models demonstrating error around 15%–30% and imprecision (CV_{Sim} of 30%–40%). Model-estimated coefficients of variation (CV_{Mod}) generally followed levels of precision in the associated parameters (Fig. 6).

Fig. 4. Relative error for estimated movement and reporting rate parameters in simulation Scenarios 1–4. Models with a frequency of convergence less than 95% are indicated by an asterisk (*). Model codes are provided in Table 1, but are grouped by model type and not scenario. Associated simulation model settings are given in parentheses. CC, Cape Cod – Gulf of Maine; GB, Georges Bank; SN, Southern New England – Mid-Atlantic.



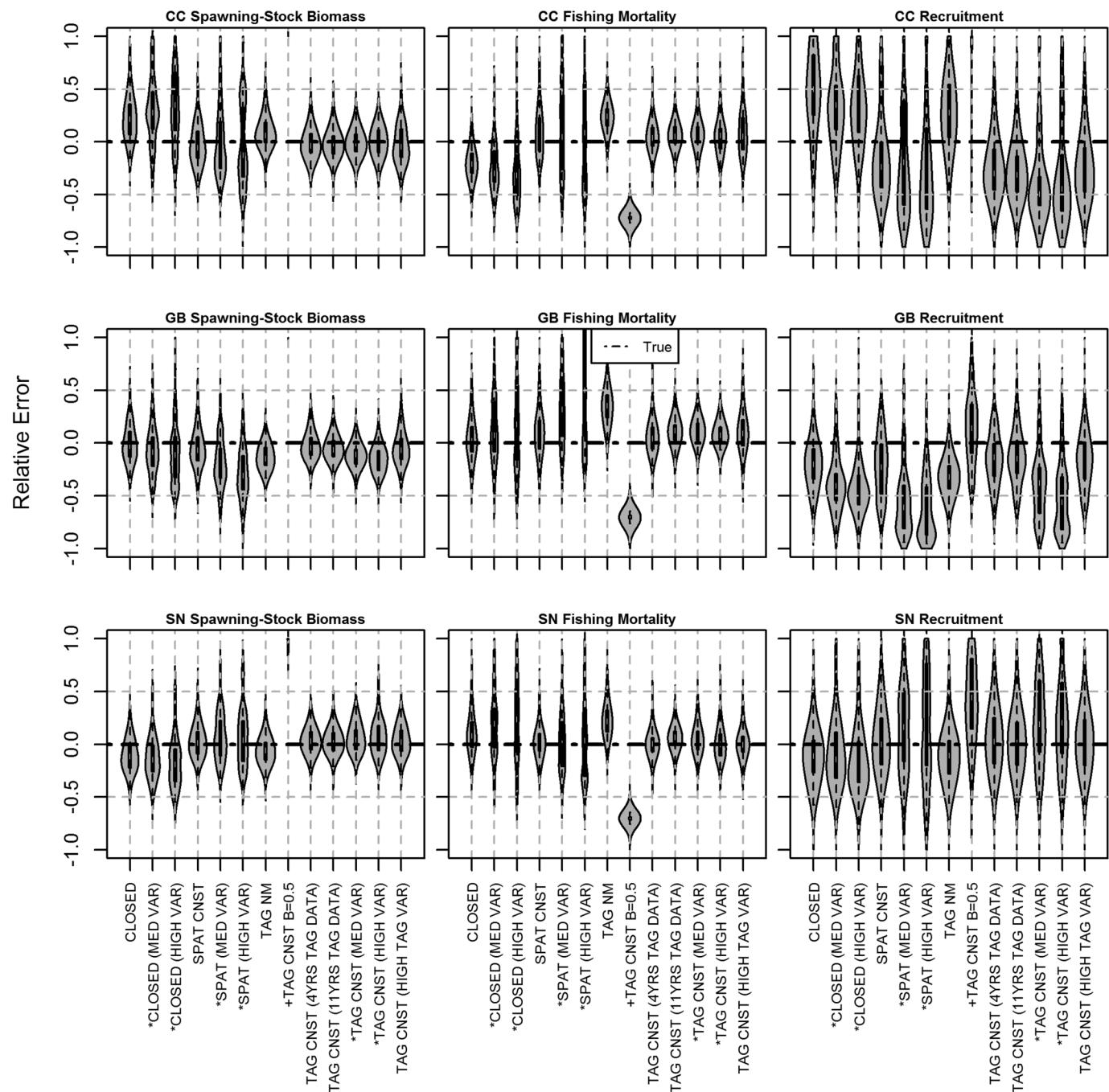
Supplementary Material Figs. S1–S9¹ illustrated the results of the CLOSED and TAG CNST base models when simulated residency rates were around 60%, 80%, and 100%. These results clearly demonstrated the robustness of the tag-integrated model to varying levels of exchange among populations, while illustrating the growing problems recreating the population dynamics (especially recruitment; Fig. S3¹) that the closed population model had as movement rates increased.

The decrease in the number of years of tagging data from 11 to 4 years (Scenario 2) decreased precision of parameter estimates in the tag-integrated model (TAG CNST; CV_{Sim} increased on average

by 2 percentage points), but error levels were unaffected (Figs. 3–5 and Tables 5–6). Reporting rate and movement estimates particularly demonstrated decreases in precision (CV_{Sim} for movement rates doubled), while model-estimated coefficients of variation increased (CV_{Mod} increased by 3–4 percentage points; Fig. 6). The inclusion of tagging data impacted model performance immediately, with precision increasing in the first year of releases, especially for estimates of fishing mortality (Supplementary Fig. S10¹).

Across all scenarios, the spatially explicit tag-integrated model (TAG CNST) was able to accurately estimate reporting rate (mean percent error less than 5%). Fixing reporting rate at the wrong

Fig. 5. Relative error for terminal year spawning-stock biomass, fishing mortality, and recruitment estimates in simulation Scenarios 1–4. Models with a frequency of convergence less than 95% are indicated by an asterisk (*). A “+” symbol is used to indicate models with a median error greater than 1.0 in at least one of the graphed parameters. Model codes are provided in Table 1, but are grouped by model type and not scenario. Associated simulation model settings are given in parentheses. CC, Cape Cod – Gulf of Maine; GB, Georges Bank; SN, Southern New England – Mid-Atlantic.



value of 0.5 (TAG CNST $\beta = 0.5$) in Scenario 3 led to increased error levels (Figs. 3–5). However, performance did not degrade until the onset of tagging and had limited impact on movement rate estimates (Fig. 4). The interaction between reporting rate and fishing mortality led to underestimation of terminal year fishing mortality (70% mean percent error), which resulted in overestimation (greater than 100% mean percent error) of terminal year spawning stock biomass (Fig. 5 and Table 5). Model-estimated coefficients of variation for the fixed reporting rate run were similar to those from the base tag-integrated model (Fig. 6). The low model-

estimated CV_{Mod} were likely an artifact of the model estimating fewer parameters that were important for the scaling of fishing mortality estimates (i.e., due to the inverse relationship between reporting rate and fishing mortality) and did not reflect the high resulting parameter error.

As data uncertainty increased (Scenario 4), the frequency of convergence decreased for all models (except for the run with increased tagging data variance; Table 4). Therefore, model performance is worse than indicated solely by error and precision of converged estimates, and results should be interpreted carefully.

Table 5. Terminal year mean percent error of model parameter estimates for all converged runs (out of 100) for a given estimation model within a simulated scenario.

Simulation scenario	Simulation setting	Estimation model	CC			GB			SN			CC β	T_{CC-GB}	T_{CC-SN}
			SSB	F	R	SSB	F	R	SSB	F	R			
Base model (Scenario 1)	Base model	CLOSED	+23	-20	+55	-1	+4	-20	-10	+10	-11	NA	NA	NA
		TAG NM	+9	+24	+31	-13	+35	-32	-6	+23	-11	-5	NA	NA
		SPAT CNST	-3	+9	-13	-5	+11	-22	+2	+2	+4	NA	-5	+2
		TAG CNST	-2	+6	-28	-5	+11	-14	+3	+6	+23	-4	-8	-5
Tagging methodology (Scenario 2)	4 years of tagging data	TAG CNST	-2	+5	-25	-3	+5	-15	+3	0	+5	-4	-6	-5
Reporting rate treatment (Scenario 3)	Base model	TAG CNST $\beta = 0.5$	+220	-72	+298	+169	-71	+15	+143	-70	+51	NA	+19	+2
Data uncertainty (Scenario 4)	Medium variance	CLOSED*	+30	-23	+38	-7	+15	-39	-13	+18	-9	NA	NA	NA
		SPAT CNST*	+3	+11	+13	-19	+47	-55	+6	+2	+27	NA	+29	-6
		TAG CNST*	-1	+7	-37	-12	+12	-40	+6	+4	+33	-4	-3	-7
		CLOSED*	+42	-28	+40	-8	+28	-38	-17	+29	-16	NA	NA	NA
	High variance	SPAT CNST*	0	+26	-2	-29	+73	-62	+5	+22	+33	NA	+81	-30
		TAG CNST*	0	+5	-25	-15	+8	-52	+7	0	+33	-3	-2	-5
		TAG CNST	-3	+8	-24	-5	+11	-16	3	+1	+3	-4	-12	+4
		TAG 11TB*	-3	+7	-35	-4	+11	-13	+1	+8	+1	-1	-8	-11
Movement parametrization (Scenario 5)	Base model	TAG RW*	-1	+6	-38	-5	+13	-16	+6	+5	+3	-1	-12	-9
		CLOSED	+22	-19	+42	+1	+1	-4	-12	+1	-27	NA	NA	NA
		TAG NM	+9	+10	+26	-11	+23	-17	-6	+21	-24	-3	NA	NA
		SPAT CNST*	+66	-37	+132	-1	+3	-29	-41	+148	-55	NA	+341	+306
		SPAT 11TB	-6	+15	-12	-5	+11	-20	+5	+2	-1	NA	+13	-2
		SPAT RW*	-13	+30	-28	-4	+10	-20	+11	-1	+5	NA	+369	-58
		SPAT YRLY*	-32	+80	-42	+7	0	-6	+10	-2	+2	NA	+456	+76
		TAG CNST	+1	+4	+3	+8	+3	+20	-11	+19	-46	-1	-7	+36
		TAG 11TB*	-4	+6	-18	-4	+11	-15	+4	+6	0	-1	-59	-57
		TAG RW*	-4	+8	-19	-7	+13	-23	+11	+2	+4	-1	-41	-33
		TAG YRLY*	-8	+11	-34	-1	+8	-12	+4	+7	+2	-2	-60	-67

Note: An asterisk (*) indicates convergence rates were below 95% for that estimation model. CC, Cape Cod – Gulf of Maine; GB, Georges Bank; SN, Southern New England – Mid-Atlantic. SSB, spawning-stock biomass; F, fishing mortality; R, recruitment; β , reporting rate; T, movement rate; NA, not applicable.

With medium levels of data uncertainty, model precision decreased (on average a 5–10 percentage point increase in CV_{Sim}) and caused an increase in mean percent error (5–10 percentage points) for each of the estimation models (Figs. 3–5 and Tables 5–6). The spatially explicit tag-integrated model (TAG CNST) was relatively robust to model error with no appreciable changes in model accuracy. The exception was in recruitment estimates, which showed a trend of increasing error and imprecision (10–15 percentage point increases in mean error and a 10–30 percentage point decrease in precision for the tag-integrated model) as data uncertainty increased. Model-estimated coefficients of variation showed approximately proportional increases to the increase in data uncertainty, but the tag-integrated model was not as severely impacted as other estimation models (Fig. 6). The runs with high data variance demonstrated a continuing trend towards lower precision and higher error. Decreasing the effective sample size of simulated tagging data led to similar increases in imprecision for the tag-integrated models as medium variance runs.

When density-dependent movement was simulated (Scenario 5), the closed population model (CLOSED) performed similarly to when simulated movement was constant. The spatially explicit tag-integrated model (TAG CNST) also demonstrated error and precision levels comparable to when simulated movement was constant, except in recruitment estimates, which tended to have more error (e.g., Southern New England – Mid-Atlantic terminal year recruitment mean percent error changed from +23% when simulated movement was constant to -46% when it was density-dependent; Figs. 7–8 and Tables 5–6). Also, some movement estimates demonstrated higher error with decreased precision (e.g., Cape Cod – Gulf of Maine to Southern New England – Mid-Atlantic movement had a mean percent error of +36% in the terminal year

and a coefficient of variation (CV_{Sim}) of 15%). The spatial model without tagging data (SPAT CNST), however, performed worse when simulated movement was density-dependent (e.g., terminal year spawning-stock biomass for the Cape Cod – Gulf of Maine stock had a mean percent error of +66%; Table 5 and Fig. 8). Additionally, convergence rates were extremely low for all of the spatial models without tagging data (except SPAT 11TB) when simulated movement was time-varying. Low convergence indicated a high degree of model instability, which caused results to not be representative (Table 4).

The tag-integrated model with 11 time blocks of movement (TAG 11TB) showed lower error (mean percent error around 5%) and more precision (CV_{Sim} of 10%) for most model parameters compared with the tag-integrated model with constant movement (TAG CNST) regardless of simulated movement scenarios. The exception was for Cape Cod – Gulf of Maine movement estimates with simulated density-dependent movement (terminal year mean percent error around 60%) and movement estimates prior to the availability of tagging data (Figs. 7–9). Other data sources (e.g., catch data) were able to help inform movement estimates, but only when large year-classes moved among stocks. The cross-stock movement of the 1987 Southern New England – Mid-Atlantic year-class (age-1 in 1988) was apparent in the 1989–1990 movement parameter estimates (Fig. 9). Other tag-integrated models that estimated time-varying movement (e.g., TAG RW and TAG YRLY) demonstrated results similar to the 11 time-block movement model, but showed low convergence rates, and results were again not considered representative.

Ignorance of time-varying movement led the constant movement and closed population models to compensate for the mispecification of connectivity by adjusting recruitment estimates

Table 6. Terminal year precision (coefficient of variation, CV_{Sim}) of model parameter estimates for all converged runs (out of 100) for a given estimation model within a simulated scenario.

Simulation scenario	Simulation setting	Estimation model	CC			GB			SN			$CC\beta$	T_{CC-GB}	T_{CC-SN}
			SSB	F	R	SSB	F	R	SSB	F	R			
Base model (Scenario 1)	Base model	CLOSED	18	18	30	17	19	29	17	18	27	NA	NA	NA
		TAG NM	13	10	29	13	12	26	13	13	27	2	NA	NA
		SPAT CNST	19	21	53	19	21	38	14	14	30	NA	61	35
		TAG CNST	12	10	37	11	10	29	11	9	28	2	9	8
Tagging methodology (Scenario 2)	4 years of tagging data	TAG CNST	14	12	37	13	13	29	11	11	28	4	18	17
Reporting rate treatment (Scenario 3)	Base model	TAG CNST $\beta = 0.5$	8	8	34	7	8	30	8	8	32	NA	7	7
Data uncertainty (Scenario 4)	Medium variance	CLOSED*	24	30	31	24	29	41	23	26	34	NA	NA	NA
		SPAT CNST*	37	39	87	28	44	70	21	26	47	NA	81	102
		TAG CNST*	12	13	65	11	9	54	13	13	45	3	17	12
		CLOSED*	34	44	33	36	45	87	30	41	36	NA	NA	NA
	High variance	SPAT CNST*	46	67	98	34	56	79	26	111	55	NA	79	148
		TAG CNST*	13	15	82	14	8	68	15	15	47	4	15	12
		TAG CNST	18	19	41	16	17	32	13	13	27	6	28	19
		TAG 11TB*	12	10	43	13	12	31	15	13	28	2	29	32
Movement parametrization (Scenario 5)	Density-dependent movement	TAG RW*	13	11	49	14	12	37	15	13	35	2	40	31
		CLOSED	17	17	31	17	18	30	16	17	27	NA	NA	NA
		TAG NM	12	10	29	13	12	28	13	12	27	2	NA	NA
		SPAT CNST*	22	25	52	22	30	64	35	44	65	NA	147	79
		SPAT 11TB	25	25	65	21	23	37	21	22	31	NA	212	181
		SPAT RW*	30	29	47	19	18	34	21	18	32	NA	64	236
		SPAT YRLY*	27	26	83	22	25	41	19	20	28	NA	109	192
		TAG CNST	13	10	31	14	12	28	10	10	31	2	22	15
		TAG 11TB*	13	10	33	13	11	27	14	13	30	2	94	114
		TAG RW*	14	12	30	11	10	27	14	12	29	2	39	55
		TAG YRLY*	13	10	34	13	12	29	16	14	30	2	137	177

Note: These values are not model-estimated coefficients of variation (i.e., CV_{Mod}), but a measure of the range for a given parameter estimate across runs. Values are in percentages. An asterisk (*) indicates convergence rates were below 95% for that estimation model. CC, Cape Cod – Gulf of Maine; GB, Georges Bank; SN, Southern New England – Mid-Atlantic; SSB, spawning-stock biomass; F, fishing mortality; R, recruitment; β , reporting rate; T, movement rate; NA, not applicable.

(Fig. 10). Because these estimation models inherently had a mismatch between the true and estimated net immigration rates, the observed catch and survey data contained signals that the models had trouble rectifying. Both the closed population and constant movement models attempted to account for the discrepancy by modifying the number of “new” fish in each stock area. The only available means to do so (for models that were not able to alter movement rates) was by changing recruitment estimates. Therefore, there was a trade-off between movement rates and recruitment estimates resulting from process error in movement parametrizations, which led to increased error, especially in historical recruitment estimates. On the other hand, when movement was allowed to change over time (e.g., the 11 time-block movement parametrization), the tag-integrated model was better able to recreate movement trends and accurately estimate historical recruitment (Fig. 9).

Discussion

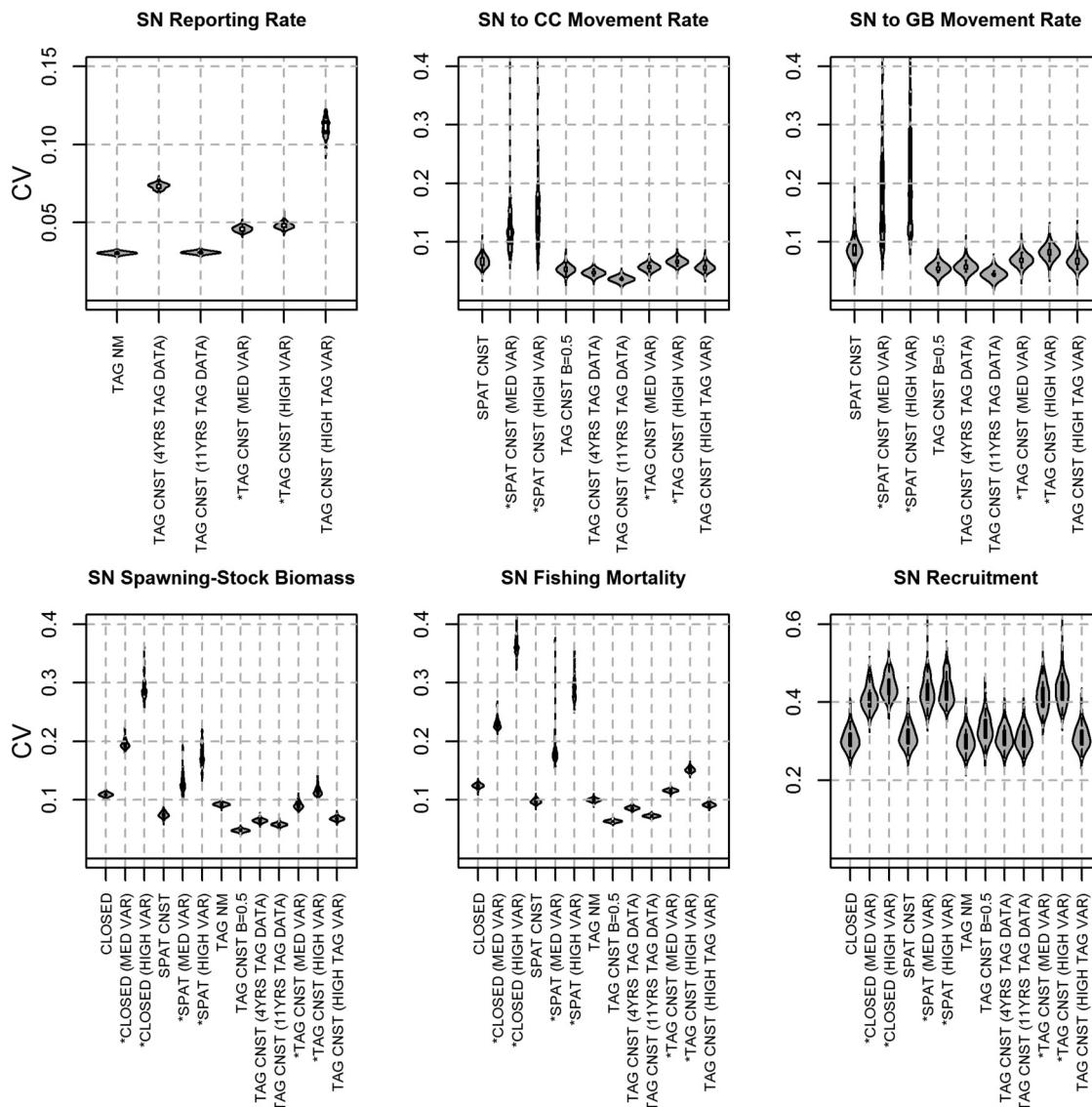
Because the tag-integrated framework is relatively new, the performance of this type of spatially explicit modelling approach is not yet thoroughly understood. Applications of tag-integrated models often fall short of expectations from simulations because of difficulty replicating the fine-scale spatial patterns or uncertainty inherent in real world data and population processes (Maunder 1998, 2001). However, our results provide a strong complement to previous simulations with tag-integrated modeling approaches (e.g., Maunder 1998, 2001; Hulson et al. 2011, 2013) and together with these prior studies suggest a number of important generalizations. Perhaps the most surprising result was the relatively robust performance of the closed population models when

confronted with high population connectivity. Although tag-integrated models were more accurate and precise, the increased data demands and complexity may limit their usefulness except for data-rich species. However, tag-integrated models provide a unique opportunity to develop spatially explicit operating models within a unified framework that can inform simulation parameters directly from observed data. By using the estimation-simulation framework where the tag-integrated model is first fit to observed data and then used as an operating model (e.g., the use of Taylor et al.’s (2011) tag-integrated model results by Kerr et al. (2013)), it is possible to develop spatial operating models that can evaluate the ability of more conventional stock assessment techniques to achieve management goals.

Our simulations also provided a more detailed understanding of metapopulation structure, especially how population dynamics models interpret spatial processes. Similar to the simulations by Hart and Cadrin (2004), our results suggest that the smallest component of the metapopulation is the most influenced by connectivity and may require careful monitoring and management to avoid overfishing. For instance, when connectivity was ignored in the closed population models, spawning-stock biomass for the Cape Cod – Gulf of Maine stock, the smallest unit in our simulations, was consistently overestimated, while fishing mortality was underestimated.

Misinterpretation of the population dynamics was not limited to the closed population models. A major impediment for interpreting spatially explicit population dynamics is attempting to decipher the interplay of movement and recruitment. Closed population models are able to account for fluctuations in abundance mainly through adjustment of recruitment estimates, while spa-

Fig. 6. Model-estimated coefficients of variation (CV_{Mod}) for terminal year parameter estimates in simulation Scenarios 1–4. The CV is not the calculated precision in parameter estimates within a given simulation scenario (i.e., CV_{Sim}), but the actual model-based coefficient of variation (CV_{Mod}) calculated from the inversion of the Hessian matrix in AD Model Builder for each model run. Models with a frequency of convergence less than 95% are indicated by an asterisk (*). Model codes are provided in Table 1, but are grouped by model type and not scenario. Associated simulation model settings are given in parentheses. Note the difference in scales along the y axis. CC, Cape Cod – Gulf of Maine; GB, Georges Bank; SN, Southern New England – Mid-Atlantic.



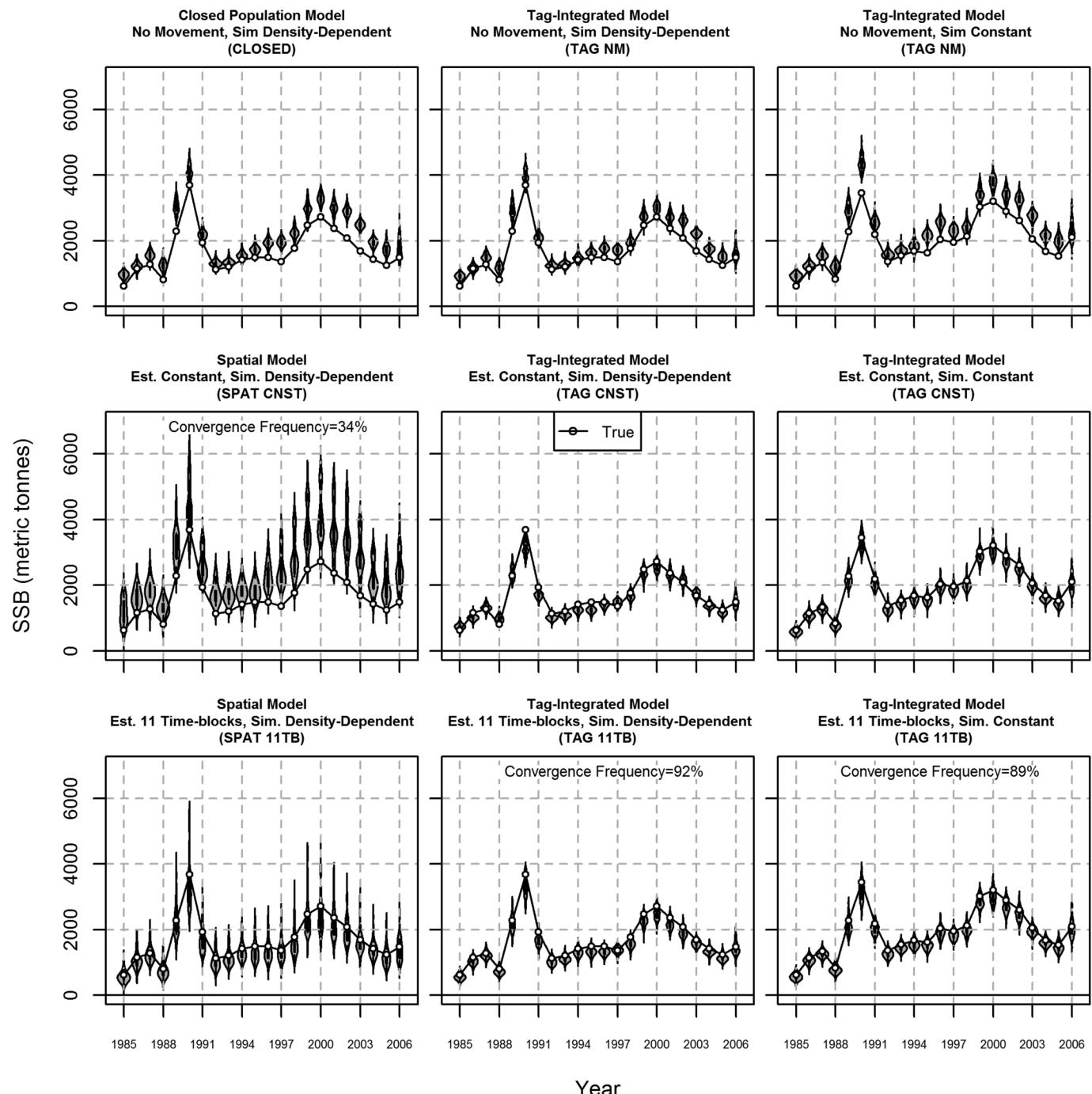
tially explicit approaches can do so by changing either movement or recruitment values. Both approaches may lead to inaccurate recruitment estimates (error is usually larger in closed population models in the presence of connectivity), and correlation of movement and recruitment parameters is common in spatial models (Maunder 1998; Maunder and Punt 2013). It is expected that most assessment models suffer from some degree of process error caused by misspecifying movement dynamics, and no single technique can be prescribed to solve the problem. Before performing any stock assessment, a thorough evaluation of connectivity pathways and spatial dynamics of the application species should be undertaken, while stationarity assumptions in movement parametrizations of spatially explicit models must be carefully evaluated. Developing a spatial operating model and testing the performance of a particular assessment across a range of likely movement scenarios is one of the few ways to determine whether

a given approach is robust to process error caused by misspecified connectivity assumptions.

Despite the difficulties in dealing with spatial data, tagging information is often unwarrantedly overlooked in stock assessment applications, considering it can improve estimates of fishing mortality (Maunder and Punt 2013). One reason that tagging data has been underutilized is the difficulty in estimating reporting rate. In previous tag-integrated studies, the impact of reporting rate on model estimates has either been ignored by assuming 100% detection of tags (e.g., Maunder 2001; Hulson et al. 2011, 2013) or by placing strong penalties on estimates (e.g., Taylor et al. 2011). If reporting rate is not known accurately, fixing it can lead to large error in tag-integrated model estimates.

Alternatively, our study illustrates that tag-integrated models are capable of estimating reporting rate when it is time-invariant. Accurate and uncorrelated reporting rate estimates have only

Fig. 7. Time-series estimates of Cape Cod – Gulf of Maine spawning-stock biomass for Scenario 5. True values are given by the black line with white dots. All models had a frequency of convergence greater than 95% unless otherwise noted on the plot. Model codes are provided in Table 1.

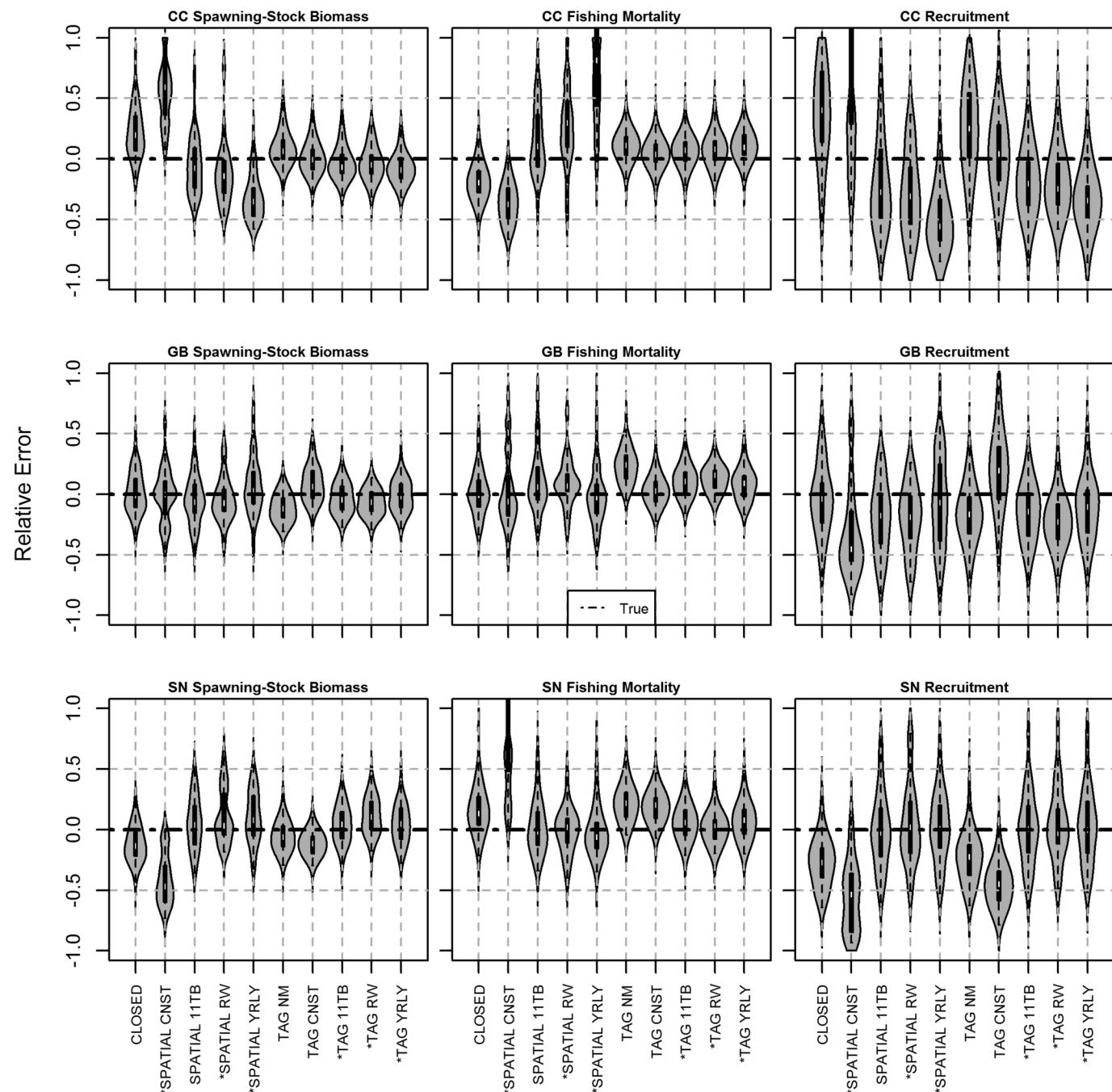


been achieved in a few fisheries models (Hoenig et al. 1998; Eveson et al. 2007). However, tagging models often attempt to estimate natural mortality along with reporting rate and fishing mortality, whereas perfect knowledge of natural mortality was assumed in this study. Considering that reporting rates are dependent on fishermen's behavior and outreach efforts, it is unlikely that they are constant over time. Nonstationarity in reporting rates will lead to error in tag-integrated models that assume time-invariant reporting rate parametrizations.

A promising approach to improve reporting rate estimates is the incorporation of data on, or estimates of, reporting rate (e.g., observer recaptures or high reward tag ratio estimates) directly

within the likelihood function by including a term for fitting estimated reporting rates to observations (Polacheck et al. 2006; Eveson et al. 2007). The power of integrated analysis lies in the ability to maintain consistent modeling assumptions and parameter estimates across submodels (Maunder and Punt 2013). When reporting rate is estimated within the stock assessment, especially if direct observations are included in the objective function, results are likely to be improved compared with treating reporting rate as an input parameter. It is unlikely that estimates of reporting rate using observed data will be as accurate as our results indicate (particularly when nonstationarity occurs), but we

Fig. 8. Relative error for terminal year spawning-stock biomass, fishing mortality, and recruitment estimates in simulation Scenario 5 with simulated density-dependent movement. Models with a frequency of convergence less than 95% are indicated by an asterisk (*). Model codes are provided in Table 1. CC, Cape Cod – Gulf of Maine; GB, Georges Bank; SN, Southern New England – Mid-Atlantic.



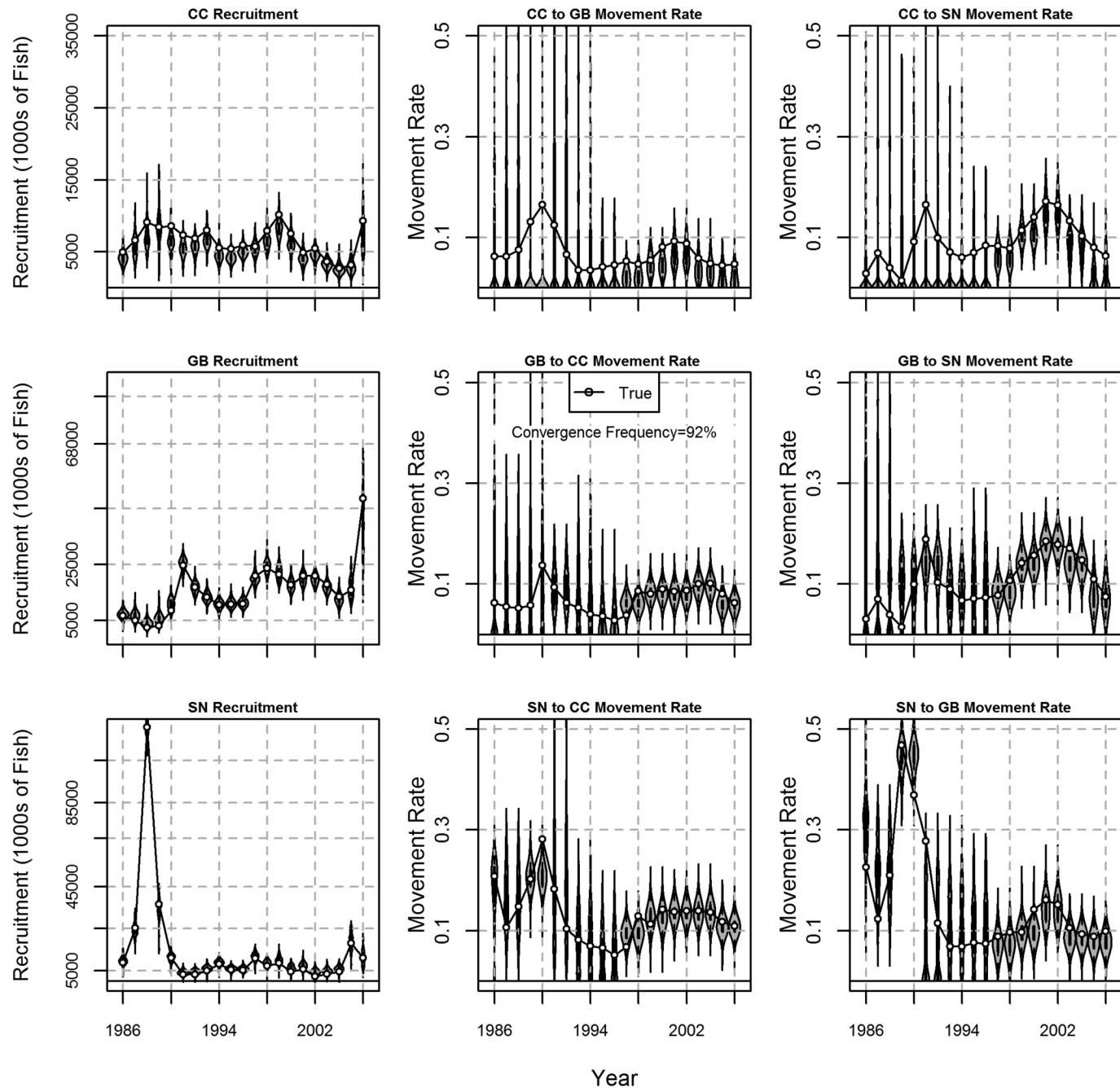
suggest that tag-integrated models are better equipped to estimate reporting rate than previously recognized.

The inclusion of tagging data also helped to reduce error and model-estimated coefficients of variation, while increasing precision in all base tag-integrated models. Our results support the findings of Hulson et al. (2013) that the decline in parameter uncertainty gained from incorporating tagging data in an assessment could lead to higher total allowable catches (TACs). The reason for the increase in TACs is because United States annual catch limits (ACLs) should reflect the level of assessment uncertainty, and a decline in the scientific uncertainty buffer would result from increases in assessment precision. Additionally, we

demonstrated that model-estimated uncertainty (CV_{Mod}) closely reflects precision estimates (CV_{Sim}) across runs within a given simulation scenario. The overlap of uncertainty estimates suggests that model-estimated uncertainty could be reliably used in the determination of ACLs, but our simulations do not take into account major sources of model misspecification error that could cause underestimation of uncertainty (e.g., if natural mortality is fixed at an incorrect value).

An interesting aspect of Scenario 2 was the immediate improvement in precision when tagging data became available within the tag-integrated model. Considering that the length of the tagging time series had limited impact on error and precision, even short,

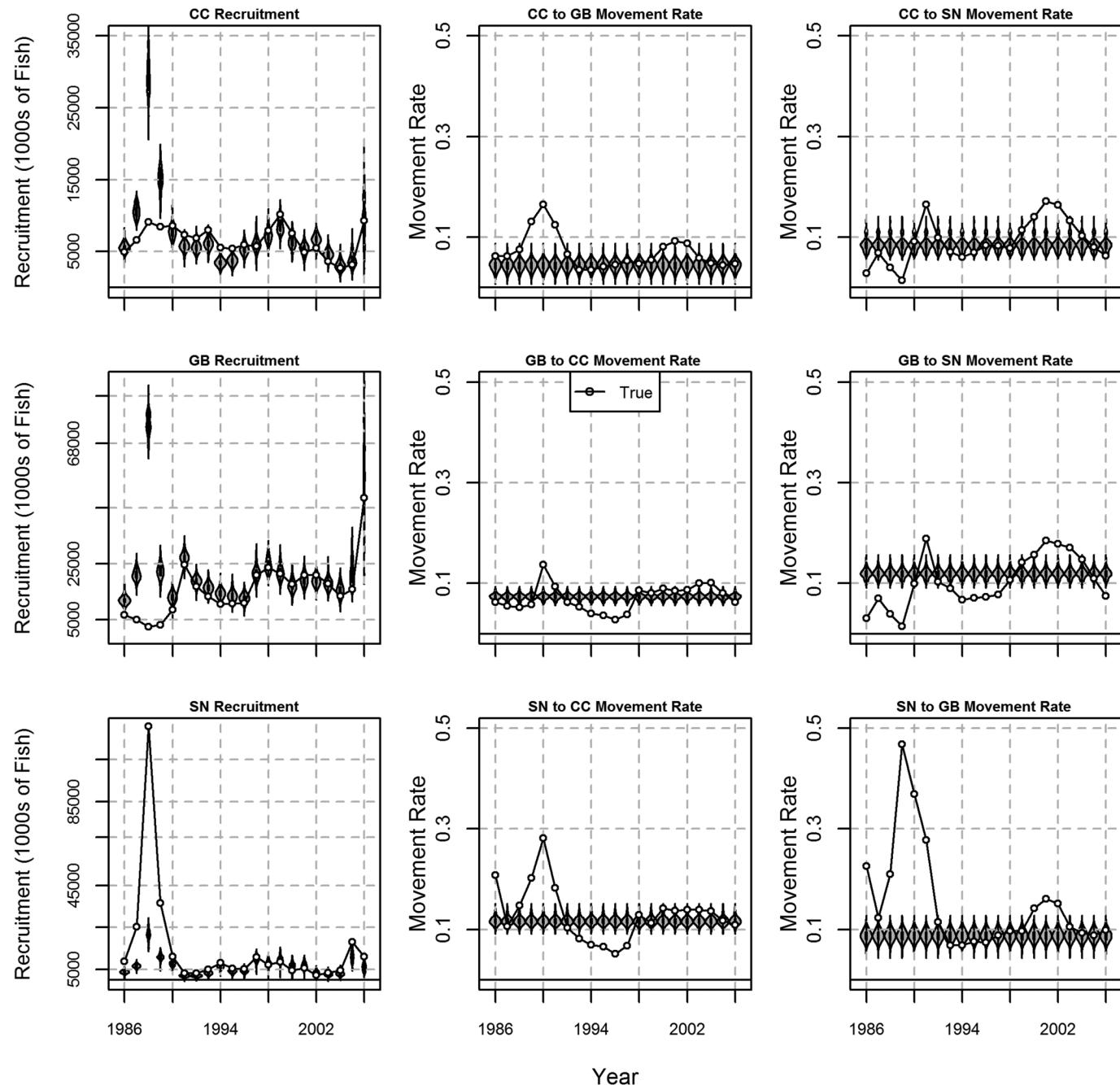
Fig. 9. Time-series of recruitment and movement rate estimates for the tag-integrated model assuming 11 time blocks of movement. Results are from Scenario 5 with simulated density-dependent movement. True values are given by the black line with white dots. Note the difference in scales along the y axis. CC, Cape Cod – Gulf of Maine; GB, Georges Bank; SN, Southern New England – Mid-Atlantic.



well-designed tagging studies may be sufficient to improve stock assessments. However, the representativeness of the tagging data for the entire time series of the model is critical for short tagging studies (Maunder 1998). There are many practical difficulties in obtaining reliable tagging data, but as tagging technology improves and becomes less expensive, the ability to incorporate different types of tags (e.g., satellite and archival tags) directly into the objective function of tag-integrated models (e.g., Taylor et al. 2011) may help improve model performance by allowing researchers to better determine the representativeness of individual tag tracks (see Sippel et al. 2014 for a review of the potential use of electronic tags in stock assessments).

As demonstrated in Scenario 5, if an incorrect movement parameterization is implemented based on erroneous indications of time-invariant movement rates, high error in model outputs will result. A balance between the findings of Scenario 2 and Scenario 5 suggests a benefit to developing biennial or triennial tag release programs with continual recaptures. A tagging study with periodic releases would provide continual feedback to the tag-integrated assessment on fishing mortality and changes in connectivity patterns while improving the precision of estimates and providing a more cost-effective alternative to yearly tagging. As with any long-term tagging program, careful consideration of changes in reporting rate over time would be required.

Fig. 10. Time-series of recruitment and movement rate estimates for the tag-integrated model assuming constant movement. Results are from Scenario 5 with simulated density-dependent movement. True values are given by the black line with white dots. Note the difference in scales along the y axis. CC, Cape Cod – Gulf of Maine; GB, Georges Bank; SN, Southern New England – Mid-Atlantic.



Our simulations indicate that attempting to model metapopulation dynamics without tagging information is extremely difficult. Contrary to the findings of Hulson et al. (2011), in the simulations performed here, the spatial model not informed by tagging data was relatively imprecise. Although other data sources included signals that helped inform movement (e.g., catch data), these signals were normally weak.

Increases in data uncertainty were detrimental to the performance of all model types in the simulated system. Correlation analysis demonstrated that selectivity parameters tended to become highly confounded when uncertainty in either catch or survey age data was increased. The quality of age composition data is critical for all age-structured stock assessment models (Crone

et al. 2013). The tag-integrated model performed better than the closed population model when poor age sampling was present, but these results were reliant on low error in tagging data. Future work with tag-integrated models should investigate alternative functional forms for selectivity that may be more robust (e.g., Hulson and Hanselman 2014). Tagging data that includes age information may also be useful in improving estimates of fishery selectivity and for estimating age-dependent movement, but was unavailable for the yellowtail flounder tagging study and was not simulated here.

We believe our results are generally applicable, especially for flounder-like species that demonstrate metapopulation structure, but the inherent limitations of simulation studies should be

closely considered before attempting to broadly apply these findings without further research. Including new data in an assessment (e.g., tagging data) does not always translate into more accurate outputs if different signals are present in the various data sources (Chen et al. 2003). However, our simulations and those of Hulson et al. (2011, 2013) indicate that tag-integrated models perform comparably, if not better than, closed population models in many situations, particularly if representative tagging data can be obtained. We suggest that when reliable tagging data are available, both closed population and spatially explicit tag-integrated approaches should be run simultaneously for an extended time period (e.g., 5–10 years). By running both models, it allows a better determination of robustness and stability of the respective frameworks and could be useful in instances where the performance of a model used for the basis of management advice deteriorates. Additionally, spatially explicit operating models such as ours could be used to develop biological reference points that account for complex population structure (e.g., Ying et al. 2011) and help improve the understanding of how spatial dynamics may influence regional management and the setting of appropriate biomass targets for interconnected stocks.

Acknowledgements

We thank Paul Rago, Terry Quinn, and two anonymous reviewers for their thorough review of the manuscript. D.R.G. also thanks his graduate committee, including Brian Rothschild and Geoff Cowles, for providing feedback and guidance. Funding for this research was provided by the Massachusetts Marine Fisheries Institute and a NOAA–Sea Grant Population Dynamics fellowship.

References

- Alade, O.A. 2008. A simulation-based approach for evaluating the performance of a yellowtail flounder (*Limanda ferruginea*) movement-mortality model. Ph.D. Dissertation, University of Maryland Eastern Shore.
- Berger, A.M., Jones, M.L., Zhao, Y., and Bence, J.R. 2012. Accounting for spatial population structure at scales relevant to life history improves stock assessment: the case for Lake Erie walleye *Sander vitreus*. *Fish. Res.* **115–116**: 44–59. doi:[10.1016/j.fishres.2011.11.006](https://doi.org/10.1016/j.fishres.2011.11.006).
- Cadrin, S.X. 2006. Yellowtail flounder tagging study. Northeast Consortium Interim Final Report.
- Cadrin, S.X. 2010. Interdisciplinary analysis of yellowtail flounder stock structure off New England. *Rev. Fish. Sci.* **18**: 281–299. doi:[10.1080/10641262.2010.506251](https://doi.org/10.1080/10641262.2010.506251).
- Cadrin, S.X., and Secor, D. 2009. Accounting for spatial population structure in stock assessment: past, present, and future. In *The future of fisheries science in North America*. Edited by R.J. Beamish and B.J. Rothschild. Springer Publishing, pp. 405–426.
- Chen, Y., Chen, L., and Stergiou, K.I. 2003. Impacts of data quantity on fisheries stock assessment. *Aquat. Sci.* **65**: 92–98. doi:[10.1007/s000270300008](https://doi.org/10.1007/s000270300008).
- Cope, J.M., and Punt, A.E. 2011. Reconciling stock assessment and management scales under conditions of spatially varying catch histories. *Fish. Res.* **107**: 22–38. doi:[10.1016/j.fishres.2010.10.002](https://doi.org/10.1016/j.fishres.2010.10.002).
- Crone, P., Maunder, M.N., Valero, J., McDaniel, J., and Seemens, B. 2013. Selectivity: theory, estimation, and application in fishery stock assessment models [online]. CAPAM Workshop Series Rep. 1. Available from http://www.capamresearch.org/sites/default/files/capamresearch.org/sites/workshops/selectivity/CAPAM_Selectivity%20Workshop_Series%20Report_August%202013.pdf [accessed 17 November 2014].
- Deroba, J.J., Butterworth, D.S., Methot, R.D., De Oliveira, J.A.A., Fernandez, C., Nielsen, A., Cadrin, S.X., Dickey-Collas, M., Legault, C.M., Ianelli, J., Valero, J.L., Needle, C.L., O’Malley, J.M., Chang, Y.-J., Thompson, G.G., Canales, C., Swain, D.P., Miller, D.C.M., Hintzen, N.T., Bertignac, M., Ibaibarriaga, L., Silva, A., Murta, A., Kell, L.T., de Moor, C.L., Parma, A.M., Dichmont, C.M., Restrepo, V.R., Ye, Y., Jardim, E., Spencer, P.D., Hanselman, D.H., Blaylock, J., Mood, M., and Hulson, P.-J.F. 2015. Simulation testing the robustness of stock assessment models to error: some results from the ICES strategic initiative on stock assessment methods. *ICES J. Mar. Sci.* doi:[10.1093/icesjms/fst237](https://doi.org/10.1093/icesjms/fst237).
- Eveson, J.P., Polacheck, T., and Laslett, G.M. 2007. Incorporating fishery observer data into an integrated catch-at-age and multiyear tagging model for estimating mortality rates and abundance [online]. *Fish. Bull.* **105**: 493–508. Available from <http://fishbull.noaa.gov/1054/eveson.pdf> [accessed 17 November 2014].
- Fournier, D.A., Sibert, J.R., Majkowski, J., and Hampton, J. 1990. MULTIFAN: a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). *Can. J. Fish. Aquat. Sci.* **47**(2): 301–317. doi:[10.1139/f90-032](https://doi.org/10.1139/f90-032).
- Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., and Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim. Method. Softw.* **27**: 233–249. doi:[10.1080/10556788.2011.597854](https://doi.org/10.1080/10556788.2011.597854).
- Goethel, D.R. 2014. Development and simulation testing of a spatially explicit tag-integrated model with application to yellowtail flounder off of New England. Ph.D. dissertation, University of Massachusetts–Dartmouth.
- Goethel, D.R., Quinn, T.J., II, and Cadrin, S.X. 2011. Incorporating spatial structure in stock assessment: movement modeling in marine fish population dynamics. *Rev. Fish. Sci.* **19**: 119–136. doi:[10.1080/10641262.2011.557451](https://doi.org/10.1080/10641262.2011.557451).
- Goethel, D.R., Legault, C.M., and Cadrin, S.X. 2015. Demonstration of a spatially explicit, tag-integrated stock assessment model with application to three interconnected stocks of yellowtail flounder off of New England. *ICES J. Mar. Sci.* **72**(1): 164–177. doi:[10.1093/icesjms/fsu014](https://doi.org/10.1093/icesjms/fsu014).
- Guan, W., Cao, J., Chen, Y., and Cieri, M. 2013. Impacts of population and fishery spatial structures on fishery stock assessment. *Can. J. Fish. Aquat. Sci.* **70**(8): 1178–1189. doi:[10.1139/cjfas-2012-0364](https://doi.org/10.1139/cjfas-2012-0364).
- Hampton, J., and Fournier, D.A. 2001. A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. *Mar. Freshw. Res.* **52**: 937–963. doi:[10.1071/MF01049](https://doi.org/10.1071/MF01049).
- Hart, D.R., and Cadrin, S.X. 2004. Yellowtail flounder (*Limanda ferruginea*) off the Northeastern United States: implications of movement among stocks. In *Applications in RAMAS*. Edited by H.R. Akçakaya, M. Burgman, O. Kindvall, C.C. Wood, V. Sjögren-Gulve, J.S. Hatfield, and M.A. McCarthy. Oxford University Press, New York. pp. 230–243.
- Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman and Hall, New York.
- Hoenig, J.M., Barrowman, N.J., Pollock, K.H., Brooks, E.N., Hearn, W.S., and Polacheck, T. 1998. Models for tagging data that allow incomplete mixing of newly tagged animals. *Can. J. Fish. Aquat. Sci.* **55**(6): 1477–1483. doi:[10.1139/f97-258](https://doi.org/10.1139/f97-258).
- Hulson, P.-J.F., and Hanselman, D.H. 2014. Tradeoffs between bias, robustness, and common sense when choosing selectivity forms. *Fish. Res.* **158**: 63–73. doi:[10.1016/j.fishres.2013.12.016](https://doi.org/10.1016/j.fishres.2013.12.016).
- Hulson, P.-J.F., Miller, S.E., Ianelli, J.N., and Quinn, T.J., II. 2011. Including mark-recapture data into a spatial age-structured model: walleye pollock (*Theragra chalcogramma*) in the eastern Bering Sea. *Can. J. Fish. Aquat. Sci.* **68**(9): 1625–1634. doi:[10.1139/f2011-060](https://doi.org/10.1139/f2011-060).
- Hulson, P.-J.F., Quinn, T.J., II, Hanselman, D.H., and Ianelli, J.N. 2013. Spatial modeling of Bering Sea walleye pollock with integrated age-structured assessment models in a changing environment. *Can. J. Fish. Aquat. Sci.* **70**(9): 1402–1416. doi:[10.1139/cjfas-2013-0020](https://doi.org/10.1139/cjfas-2013-0020).
- Hurtado-Ferro, F., Punt, A.E., and Hill, K.T. 2014. Use of multiple selectivity patterns as a proxy for spatial structure. *Fish. Res.* **158**: 102–115. doi:[10.1016/j.fishres.2013.10.001](https://doi.org/10.1016/j.fishres.2013.10.001).
- Kerr, L.A., and Goethel, D.R. 2014. Simulation modelling as a tool for synthesis of stock identification information. In *Stock identification methods: an overview*. Edited by S.X. Cadrin, L.A. Kerr, and S. Mariani. Elsevier Science and Technology, Burlington, Mass. pp. 501–533. doi:[10.1016/B978-0-12-397003-9.00021-7](https://doi.org/10.1016/B978-0-12-397003-9.00021-7).
- Kerr, L.A., Cadrin, S.X., Secor, D.H., and Taylor, N. 2013. A simulation tool to evaluate effects of mixing between Atlantic bluefin tuna stocks [online]. Collect. Vol. Sci. Pap. ICCAT. **69**: 742–759. Available from https://www.iccat.int/Documents/CVSP/CV069_2013/n_2/CV069020742.pdf [accessed 17 November 2014].
- Kolody, D.S., Jumpanen, P.C., Ricard, D.G., Hartog, J.R., Preece, A.L., and Polacheck, T. 2004. Extracts from SESAME: a simulation-estimation stock assessment model evaluation project focused on large pelagic species [online]. CSIRO Mar. Lab. Rep. 241. Available from http://www.spc.int/DigitalLibrary/Doc/FAME/Meetings/SCBT/17/MWG_3.pdf [accessed 17 November 2014].
- Kritzer, J.P., and Sale, P.F. 2004. Metapopulation ecology in the sea: from Levins’ model to marine ecology and fisheries science. *Fish. Fish.* **5**: 131–140. doi:[10.1111/j.1467-2979.2004.00131.x](https://doi.org/10.1111/j.1467-2979.2004.00131.x).
- Legault, C.M., Alade, L., Stone, H.H., Gross, W.E., and Stone, H.H. 2014. Stock assessment of Georges Bank yellowtail flounder for 2013. TRAC Work. Pap. 2014.
- MacCall, A.D. 1990. Dynamic geography of marine fish populations. University of Washington Press, Seattle, Wash.
- Maunder, M.N. 1998. Integration of tagging and population dynamics models in fisheries stock assessment [online]. Ph.D. dissertation, University of Washington, Seattle. Available from <http://hdl.handle.net/1773/5284> [accessed 17 November 2014].
- Maunder, M.N. 2001. Integrated tagging and catch-at-age analysis (ITCAAN): model development and simulation testing. In *Spatial processes and management of marine populations*. Edited by G.H. Kruse, N. Bez, A. Booth, M.W. Dorn, S. Hills, R.N. Lipcius, D. Pelletier, C. Roy, S.J. Smith, and D. Witherell. University of Alaska Sea Grant, Fairbanks, Alaska. AK-SG-01-02. pp. 123–146.
- Maunder, M.N., and Punt, A.E. 2013. A review of integrated analysis in fisheries stock assessment. *Fish. Res.* **142**: 61–74. doi:[10.1016/j.fishres.2012.07.025](https://doi.org/10.1016/j.fishres.2012.07.025).
- Methot, R.D., and Wetzel, C.R. 2013. Stock Synthesis: a biological and statistical

- framework for fish stock assessment and fishery management. *Fish. Res.* **142**: 86–99. doi:[10.1016/j.fishres.2012.10.012](https://doi.org/10.1016/j.fishres.2012.10.012).
- NEFSC. 2012a. Assessment or data updates of 13 Northeast groundfish stocks through 2010 [online]. US Dept. Comm., NOAA Fisheries. NEFSC Ref. Doc. 12-06. Available from <http://www.nefsc.noaa.gov/publications/crd/crd1206/1206.pdf> [accessed 17 November 2014].
- NEFSC. 2012b. 54th Northeast regional stock assessment workshop (54th SAW) assessment report [online]. US Dept. Comm., NOAA Fisheries. NEFSC Ref. Doc. 12-18. Available from <http://www.nefsc.noaa.gov/publications/crd/crd1218/crd1218.pdf> [accessed 17 November 2014].
- Polacheck, T., Eveson, J.P., Laslett, G.M., Pollock, K.H., and Hearn, W.S. 2006. Integrating catch-at-age and multiyear tagging data: a combined Brownie and Petersen estimation approach in a fishery context. *Can. J. Fish. Aquat. Sci.* **63**(3): 534–548. doi:[10.1139/f05-232](https://doi.org/10.1139/f05-232).
- Pollock, K.H., Hoenig, J.H., Hearn, W.S., and Calingaert, B. 2002. Tag reporting rate estimation: 2. Use of high-reward tagging and observers in multiple-component fisheries. *N. Am. J. Fish. Manage.* **22**: 727–736. doi:[10.1577/1548-8675\(2002\)022<0727:TRREUO>2.0.CO;2](https://doi.org/10.1577/1548-8675.2002.022<0727:TRREUO>2.0.CO;2).
- Porch, C., Kleiber, P., Turner, S., Sibert, J., Bailey, R., and Cort, J.L. 1998. The efficacy of VPA models in the presence of complicated movement patterns [online]. Collect. Vol. Sci. Pap. ICCAT. **50**: 591–622. Available from www.iccat.es/documents/cvsp/cv050_1998/no_2/ CV050020591.pdf [accessed 17 November 2014].
- Sippel, T., Evenson, J.P., Galuardi, B., Lam, C., Hoyle, S., Maunder, M.N., Kleiber, P., Carvalho, F., Tsontos, V., Teo, S.L.H., Aires-da-Silva, A., and Nicol, S. 2014. Using movement data from electronic tags in fisheries stock assessment: a review of models, technology, and experimental design. *Fish. Res.* doi:[10.1016/j.fishres.2014.04.006](https://doi.org/10.1016/j.fishres.2014.04.006).
- Smedbol, R.K., and Stephenson, R.L. 2001. The importance of managing within-species diversity in cod and herring fisheries of the North-Western Atlantic. *J. Fish. Biol.* **59**: 109–128. doi:[10.1111/j.1095-8649.2001.tb01382.x](https://doi.org/10.1111/j.1095-8649.2001.tb01382.x).
- Taylor, N.G., McAllister, M.K., Lawson, G.L., Carruthers, T., and Block, B.A. 2011. Atlantic bluefin tuna: a novel multistock spatial model for assessing population biomass. *PLoS ONE*, **6**: e27693. doi:[10.1371/journal.pone.0027693](https://doi.org/10.1371/journal.pone.0027693). PMID: 22174745.
- Wood, A.D., and Cadrian, S.X. 2013. Mortality and movement of yellowtail flounder (*Limanda ferruginea*) tagged off New England. *Fish. Bull.* **111**: 279–287. doi:[10.7755/FB.111.3.6](https://doi.org/10.7755/FB.111.3.6).
- Ying, Y., Chen, Y., Lin, L., and Gao, T. 2011. Risks of ignoring fish population spatial structure in fisheries management. *Can. J. Fish. Aquat. Sci.* **68**(12): 2101–2120. doi:[10.1139/f2011-116](https://doi.org/10.1139/f2011-116).