

The dream and the reality: meeting decision-making time frames while incorporating ecosystem and economic models into management strategy evaluation^{1,2}

Jonathan J. Deroba, Sarah K. Gaichas, Min-Yang Lee, Rachel G. Feeney, Deirdre Boelke, and Brian J. Irwin

Abstract: Atlantic herring (*Clupea harengus*) in the Northwest Atlantic have been managed with interim harvest control rules (HCRs). A stakeholder-driven management strategy evaluation (MSE) was conducted that incorporated a broad range of objectives. The MSE process was completed within 1 year. Constant catch, conditional constant catch, and a biomass-based (BB) HCR with a 15% restriction on the interannual change in the quota could achieve more stable yields than BB HCRs without such restrictions, but could not attain as high of yields and resulted in more negative outcomes for terns (*Sterna hirundo*; a predator of herring). A similar range of performance could be achieved by applying a BB HCR annually every 3 years or every 5 years. Predators (i.e., dogfish (*Squalus acanthias*), bluefin tuna (*Thunnus thynnus*), and terns) were generally insensitive to the range of HCRs. While median net revenues were sensitive to some HCRs, time series analysis suggests that most HCRs produced a stable equilibrium of net revenue. To meet management needs, some aspects of the simulations were less than might be considered scientifically ideal, but using “models of intermediate complexity” were informative for managers and formed a foundation for future improvements.

Résumé : Les harengs (*Clupea harengus*) dans le nord-ouest de l’océan Atlantique ont été gérés en utilisant des règles de contrôle des prises (RCP) provisoires. Une évaluation des stratégies de gestion (ESG) menée par les parties prenantes et visant toute une gamme d’objectifs a été effectuée. Le processus d’ESG a pris 1 an. Des politiques de prises constantes et de prises constantes conditionnelles et une RCP basée sur la biomasse (BB) prévoyant une limite de 15 % des variations interannuelles des quotas permettent l’atteinte de rendements plus stables que des RCP BB sans de telles limites, mais ne permettent pas des rendements aussi élevés et se traduisent par des résultats plus négatifs pour les sternes (*Sterna hirundo*; des prédateurs du hareng). Une fourchette semblable de performances pourrait être possible en appliquant une RCP BB annuellement, tous les 3 ans ou tous les 5 ans. Les prédateurs (c.-à-d. squales (*Squalus acanthias*), thons (*Thunnus thynnus*) et sternes) sont généralement insensibles à la fourchette de RCP. Si les recettes médianes nettes sont sensibles à certaines RCP, l’analyse de séries chronologiques porte à croire que la plupart des RCP produisent un équilibre stable des recettes nettes. Pour répondre aux besoins de gestion, certains aspects des simulations ne sont pas idéaux du point de vue scientifique, mais l’utilisation de « modèles de complexité intermédiaire » fournit de l’information utile aux gestionnaires et sert de base pour des améliorations futures. [Traduit par la Rédaction]

Introduction

Atlantic herring (hereinafter herring; *Clupea harengus*) in the Northwest Atlantic are preyed upon by fish, seabirds, and marine mammals and can account for 20%–50% of the diet of these predators (Overholtz and Link 2007; Smith and Link 2010; Curti et al. 2013). Herring are also subject to a directed fishery, mostly using midwater trawls, purse seines, and bottom trawls, that averaged 85 000 t of landings during 2008–2017. Much of the herring landed are used as bait in the relatively high-value American lobster (*Homarus americanus*) fishery. Herring life-history traits (e.g., weight-at-age) have varied through time, and their complex stock structure creates uncertainty in their assessment and manage-

ment (Northeast Fisheries Science Center 2012; Deroba 2015). Thus, herring are of broad interest in the region, but anticipating the relative performance of management strategies (e.g., harvest control rules, HCRs) in the face of uncertainties is challenging, which makes a management strategy evaluation (MSE) for herring in the region a potentially informative tool.

The federal fisheries management process in New England usually starts when the New England Fishery Management Council (NEFMC), the political body responsible for federally managed species in the Northeast US, perceives a problem causing a management goal or objective to be unmet. Managers will propose a range of potential solutions, and a technical group, typically composed of scientists and policy analysts from state agencies and the

Received 30 March 2018. Accepted 6 July 2018.

J.J. Deroba, S.K. Gaichas, and M.-Y. Lee. NOAA Fisheries, Northeast Fisheries Science Center, 166 Water St., Woods Hole MA 02543, USA.

R.G. Feeney and D. Boelke. New England Fisheries Management Council, 50 Water Street, Mill 2, Newburyport, MA 01950, USA.

B.J. Irwin. US Geological Survey, Georgia Cooperative Fish and Wildlife Research Unit, University of Georgia, 180 E. Green St., Athens, GA 30602, USA.

Corresponding author: Jonathan J. Deroba (email: Jonathan.Deroba@noaa.gov).

¹This paper is one of a pair of companion papers published in this issue of Can. J. Fish. Aquat. Sci. (Feeney et al. 2019. Can. J. Fish. Aquat. Sci. This issue. doi:[dx.doi.org/10.1139/cjfas-2018-0125](https://doi.org/10.1139/cjfas-2018-0125)).

²This article is being published as part of the special issue “Under pressure: addressing fisheries challenges with Management Strategy Evaluation” arising from two related theme sessions sponsored by the American Institute of Fishery Research Biologists at the 147th Annual Meeting of the American Fisheries Society, Tampa, Florida, USA, August 2017.

Copyright remains with the author(s) or their institution(s). Permission for reuse (free in most cases) can be obtained from [RightsLink](https://www.rightslink.com).

National Marine Fisheries Service (NMFS), will analyze these possible solutions. Council meetings are public, and stakeholder input is solicited through open comment periods and diverse advisory panels. Once the NEFMC votes on a solution, NMFS will, after verifying that it is consistent with applicable laws, translate those solutions into regulations and enforce those regulations. Management actions, particularly contentious ones, can take several years to develop and implement.

MSE uses simulation to evaluate the trade-offs resulting from alternative management options in the face of uncertainty (Punt et al. 2016a). MSEs require time, however, for stakeholder input, data collection, and model development (Butterworth 2007; Punt et al. 2016a). As such, the process can take much longer than “traditional” management time frames (Butterworth 2007). The development time is also likely to lengthen when explicit ecosystem, multispecies, or socioeconomic considerations are included because the data and modeling needs, and subsequent uncertainties, are all greater than in a single species approach. This manuscript chronicles the development of an MSE done on a truncated timetable (~12 months) required to meet management time frames. The objectives of this manuscript were to

1. Evaluate the relative performance of HCRs at meeting herring fishery objectives, including those related to predators of herring, as informed by stakeholder input, and
2. Discuss our approach to developing an MSE on a relatively truncated timetable to meet management time frames and identify the lessons learned throughout the process, especially as they relate to using MSE as a tool to advance an ecosystem-based approach to management (e.g., Plagányi et al. 2014).

During January 2016, the NEFMC requested an MSE to evaluate HCRs for herring (see Feeney et al. 2019 for more detail). Fishery managers wanted to develop a HCR that, among other things, accounted for the role of Atlantic herring as forage in the ecosystem. The exact “accounting” system was left to be defined by stakeholder-driven workshops that were open to the public, except that the main interest was in the effect that herring have on predators and not top-down effects of predators on herring. Two stakeholder workshops were conducted, one in May 2016 and another in December 2016. Members of the herring, lobster, groundfish, tuna, recreational, and whale-watching industries participated, as well as environmental nongovernment organizations (ENGO), federal and state agencies, and academics. Input from these diverse stakeholders was then utilized to construct the closed-loop simulation portion of the MSE. Notably, while the technical group knew the general scope of the modeling exercise (HCRs that account for herring as forage), detailed modeling of many components of interest could not begin in earnest until this step was finished.

Constructing an ecosystem model with an interdisciplinary team of scientists to support these management decisions would be a useful and perhaps ideal tool. However, the NEFMC desired results from the MSE within 1 year, which constrained the development of many components of the model. Deciding how to allocate scarce research effort and time to various components of model development is always challenging; the limited time available for model development made this process even more challenging. Relatively simple models linking key ecosystem components can have many advantages over more complex models for ecosystem analysis (Plagányi et al. 2014; Collie et al. 2016; Punt et al. 2016b). These “models of intermediate complexity for ecosystem assessments” (MICE; Plagányi et al. 2014) can also be a reasonable approach to conducting science to meet relatively short management timelines while ensuring the models and results reasonably describe the trade-offs among objectives and remain relevant for decision-making. The methods in this article describe the reality of what was achieved to enable managers to make informed deci-

sions within their preferred time frames. This reality is then contrasted with more comprehensive scientific approaches that could better inform future decision-making (i.e., the “dream”). This manuscript is a companion to Feeney et al. (2019), which more fully describes the management context and how this MSE was blended with fisheries management.

Methods — the reality

Herring

An MSE was developed specific to Gulf of Maine – Georges Bank Atlantic herring. The herring component was modified from Deroba (2014), and symbols are largely consistent (Table 1). The MSE was based on an age-structured simulation that considered fish from age-1 through age-8+ (age-8 and older), which is consistent with the age ranges used in the 2012 and 2015 Atlantic herring stock assessments (Northeast Fisheries Science Center 2012; Deroba 2015). The abundances at age in year one of all simulations equaled the equilibrium abundances produced by the fishing mortality rate that would reduce the population to 40% of $SSB_{F=0}$, but simulations were of sufficient length to make the starting values moot. Abundance in each subsequent age and year was calculated assuming that fish died exponentially according to an age- and year-specific total instantaneous mortality rate.

Recruitment followed Beverton–Holt dynamics (Francis 1992):

$$(1) \quad R_{1,y+1} = \frac{\left(\frac{SSB_{F=0} 1 - h}{R_{F=0} 4h}\right) SSB_y e^{\varepsilon_{R,y} - \frac{\sigma_R^2}{2}}}{1 + \left(\frac{5h - 1}{4hR_{F=0}}\right) SSB_y}$$

$$(2) \quad \varepsilon_{R,y} = \omega \varepsilon_{R,y-1} + \sqrt{1 - \omega^2} \rho_y$$

$$(3) \quad \rho \sim N(0, \sigma_R^2)$$

$$(4) \quad SSB_y = \sum_{a=1}^{8+} N_{a,y} m_a W_a$$

The variance of recruitment process errors (σ_R^2) equaled 0.36 and the degree of autocorrelation (ω) equaled 0.1, which are values consistent with recruitment estimates from a recent Atlantic herring stock assessment (Deroba 2015).

Assessment error

A stock assessment was approximated (i.e., assessment errors) similar to Punt et al. (2008) and Deroba (2014). Assessment error was modeled as a year-specific lognormal random deviation common to all ages, with first-order autocorrelation and a term that created the option to include bias ρ :

$$(5) \quad \hat{N}_{a,y} = [N_{a,y}(\rho + 1)]e^{\varepsilon_{\phi,y} - \frac{\sigma_\phi^2}{2}}$$

$$(6) \quad \varepsilon_{\phi,y} = \vartheta \varepsilon_{\phi,y-1} + \sqrt{1 - \vartheta^2} \tau_y$$

$$(7) \quad \tau \sim N(0, \sigma_\phi^2)$$

The variance of assessment errors (σ_ϕ^2) equaled 0.05 and autocorrelation (ϑ) equaled 0.7. A range of values for σ_ϕ^2 and ϑ were not evaluated because previous research using a similar approach to applying assessment errors found relative control rule performance to be robust to these quantities (Irwin et al. 2008; Punt et al. 2008; Deroba and Bence 2012). Rho (ρ) allowed for the inclusion of bias in the assessed value of abundance (see below; Deroba (2014)). Assessed spawning stock biomass (\hat{SSB}_y) was calculated similarly

Table 1. Table of symbols.

Symbol	Definition
y	Year
a	Age
R_y	Herring recruitment
SSB_y	Herring spawning stock biomass
h	Steepness
$\varepsilon_{R,y}$	Recruitment process error
σ_R^2	Variance of recruitment process error
ω	Autocorrelation of recruitment error
$N_{a,y}$	Herring abundance
m_a	Herring maturity
W_a	Herring weight
$\hat{N}_{a,y}$	Assessed herring abundance
ρ	Bias in assessed herring abundance
$\varepsilon_{\phi y}$	Assessment error
τ	Autocorrelation of assessment error
σ_ϕ^2	Variance of assessment error
B_y	Herring total biomass
\overline{SSB}_y	Assessed herring spawning stock biomass
\hat{B}_y	Assessed herring total biomass
M_a	Herring natural mortality
Y_y	Herring yield
MSY	Herring maximum sustainable yield
SSB_{up}	Upper biomass parameter for biomass-based control rule
SSB_{low}	Lower biomass parameter for biomass-based control rule
ψ	Proportion of F_{MSY} for biomass-based control rule
F_{MSY}	Fishing mortality at herring maximum sustainable yield
\bar{F}_y	Target herring fishing mortality
S_a	Herring fishery selectivity
\bar{F}_y	Fishing mortality that would remove quota from herring population
Q_y	Quota for herring fishery
F_y	Realized herring fishery mortality
$\varepsilon_{\theta,y}$	Implementation error
σ_θ^2	Variance of implementation error
P	Denotes a quantity that applies to a predator of herring
N_y^p	Predator abundance
S_y^p	Annual predator survival
R_y^p	Predator recruitment from Beverton–Holt relationship
\bar{R}_y^p	Predator recruitment as modified by a relationship with herring
v_y	Annual predator natural mortality
u	Annual predator exploitation mortality
FW_{int}	Intercept of Ford–Walford relationship
FW_{slope}	Slope of Ford–Walford relationship
B_y^p	Predator biomass
γ	Multiplier relating herring abundance threshold to predator recruitment
N_{thresh}	Threshold in herring abundance that induces a reaction from a predator
α_y^p	Intercept of Ford–Walford relationship as modified by a relationship with herring
ρ_y^p	Slope of Ford–Walford relationship as modified by a relationship with herring
δ	Exponent relating herring abundance to predator survival
\bar{W}_y	Average weight of herring
λ	Exponent relating herring average weight to tuna growth
T	Inflection point of the relationship between herring average weight and tuna growth
SSB_{MSY}	Herring spawning stock biomass at maximum sustainable yield
IAV	Interannual variation in herring yield
SSB_{MSY}^p	Predator spawning stock biomass at maximum sustainable yield
GR_y	Gross revenue
NR_y	Net operating revenues
i	Fleet (trawl, t , or purse seine, s)
q_y	Quantity landed
$c()$	Cost function
p	Function relating landings to prices
t	Denotes trawl fishery in economic model
s	Denotes purse seine fishery in economic model
a_i	Input to price and landing model
b	Input to price and landing model
γ	Input to price and landing model
α	Input to price and landing model
θ_i	Input to price and landing model
β	Input to price and landing model
ξ	Input to economic stationarity metric

Table 2. Operating model uncertainties addressed.

Operating model name	Herring productivity	Herring growth	Assessment bias
LowFastBiased	Low: high M , low h (0.44)	1976–1985: fast	60% overestimate
LowSlowBiased	Low: high M , low h (0.44)	2005–2014: slow	60% overestimate
LowFastCorrect	Low: high M , low h (0.44)	1976–1985: fast	None
LowSlowCorrect	Low: high M , low h (0.44)	2005–2014: slow	None
HighFastBiased	High: low M , high h (0.79)	1976–1985: fast	60% overestimate
HighSlowBiased	High: low M , high h (0.79)	2005–2014: slow	60% overestimate
HighFastCorrect	High: low M , high h (0.79)	1976–1985: fast	None
HighSlowCorrect	High: low M , high h (0.79)	2005–2014: slow	None

Table 3. Herring natural mortality and mean weights-at-age (kg).

Age	High M		Fast growth	Slow growth
	High M	Low M		
1	0.81	0.54	0.01	0.02
2	0.65	0.43	0.03	0.04
3	0.54	0.36	0.09	0.07
4	0.48	0.32	0.16	0.11
5	0.45	0.30	0.21	0.14
6	0.42	0.28	0.25	0.16
7	0.41	0.27	0.28	0.18
8+	0.40	0.26	0.31	0.21

to SSB_y except with $N_{a,y}$ replaced with $\hat{N}_{a,y}$, and assessed total biomass (\hat{B}_y) was calculated as the sum across ages of the product of $\hat{N}_{a,y}$ and W_a .

Operating models

The stakeholder workshops identified uncertainties about herring life-history traits and stock assessment. The effect of some of these uncertainties on HCR performance was evaluated by simulating the control rules for each of eight operating models (Table 2). The uncertainties addressed by the eight operating models included Atlantic herring natural mortality and steepness, Atlantic herring weight-at-age, and possible bias in the stock assessment beyond the unbiased measurement error ($\epsilon_{\phi y}$).

The specific values used in the operating models for each of the uncertainties were premised on data used in recent stock assessments or estimates from fits of stock assessment models (Deroba 2015). Natural mortality in recent stock assessments has varied among ages and years, with M being higher during 1996–2014 than in previous years (Northeast Fisheries Science Center 2012; Deroba 2015). Natural mortality, however, has also been identified as an uncertainty in the stock assessments, and sensitivity runs have been conducted without higher M during 1996–2014, such that M was constant among years (Northeast Fisheries Science Center 2012; Deroba 2015). To capture uncertainty in M in the MSE, operating models were run with either relatively high or low M_a (Table 3). Relatively high M_a values equaled the age-specific natural mortality rates used for the years 1996–2014 in the stock assessment. Relatively low M_a values in the MSE equaled the age-specific natural mortality rates used for the years 1965–1995 in the stock assessment. In the MSE, M_a was always time-invariant.

Uncertainty in estimates of stock-recruit parameters were represented in the MSE by using the parameters estimated by stock assessments fit with and without the higher M during 1996–2014. Stock assessment fits with higher M during 1996–2014 produced estimates of steepness and $SSB_{F=0}$ that were lower than in stock assessment fits without higher M during 1996–2014 (Tables 2 and 3). Thus, operating models with relatively high M_a always had relatively low steepness and $SSB_{F=0}$, and the opposite held with relatively low M_a (Table 2).

Table 4. Life-history traits and reference points including unfished SSB (i.e., virgin SSB), SSB_{MSY} , MSY , and F_{MSY} reference points for the Atlantic herring operating models.

Steepness	Natural		Unfished			
	mortality	Growth	SSB (t)	SSB_{MSY} (t)	MSY (t)	F_{MSY}
0.44	High	Slow	845 176	324 977	66 061	0.31
0.44	High	Fast	845 176	335 849	60 969	0.28
0.79	Low	Slow	1 347 080	369 089	129 171	0.54
0.79	Low	Fast	1 347 080	405 485	120 360	0.45

Uncertainty in Atlantic herring size-at-age was accounted for by having operating models with either fast or slow growth (i.e., weights-at-age; Table 3). Atlantic herring weight-at-age generally declined from the mid-1980s through the mid-1990s, especially at ages-3 and older, and has been relatively stable since (Deroba 2015). Reasons for the decline are speculative, and no causal relationships have been established. Likewise, the reasons why growth has declined for older ages but either has remained stable or increased at age-1 and age-2 are unclear. Thus, the terms “fast” and “slow” growth were used as generally describing the growth conditions of older Atlantic herring. Fast growth operating models had weights-at-age that equaled the 1 January weights-at-age from the most recent stock assessment averaged over the years 1976–1985, while the slow growth operating models averaged over the years 2005–2014 (Deroba 2015). In the MSE, weight-at-age was always time-invariant.

Differences in M , stock-recruit parameters, and weights-at-age led to differences in unfished (i.e., virgin) and MSY reference points among operating models (Table 4). The effect of M and stock-recruit parameters was larger than the effect of differences in weight-at-age (Table 4).

To address concerns about possible stock assessment bias, operating models with and without a positive bias were included. In operating models without bias, $\rho = 0$ and the only assessment error was that caused by the unbiased measurement errors ($\epsilon_{\phi y}$). In operating models with bias, $\rho = 0.6$, which was based on the degree of retrospective pattern in SSB from the most recent stock assessment (Deroba 2015).

Harvest control rules (HCRs)

Several classes of control rules were evaluated, including a biomass-based control rule (Katsukawa 2004), a constant catch rule, and a conditional constant catch rule (Clark and Hare 2004; Deroba and Bence 2012). The biomass-based control rule was defined by three parameters: the proportion (ψ) of F_{MSY} that dictates the maximum desired fishing mortality rate (\bar{F}), an upper SSB threshold (SSB_{up}), and a lower SSB threshold (SSB_{low}). The \bar{F} equaled the maximum when SSB was above the upper threshold, declined linearly between the upper and lower thresholds, and equaled zero below the lower threshold:

$$(8) \quad \tilde{F}_y = \begin{cases} \psi F_{MSY}, & \text{if } \overline{SSB}_y \geq SSB_{up} \\ \psi F_{MSY} \frac{\overline{SSB}_y - SSB_{low}}{SSB_{up} - SSB_{low}}, & \text{if } SSB_{low} < \overline{SSB}_y < SSB_{up} \\ 0, & \text{if } \overline{SSB}_y \leq SSB_{low} \end{cases}$$

The \tilde{F}_y was then used to set a quota in year $y + 1$. \tilde{F}_{ay} equaled \tilde{F}_y times S_a , and S_a was time- and simulation-invariant selectivity at age equal to the values for the mobile gear fishery as in Deroba (2015). \tilde{F}_y was used to set a quota in the following year to approximate the practice of using projections based on an assessment using data through year $y - 1$ to set quotas in the following year(s). Furthermore, although \tilde{F}_y was set using \overline{SSB}_y , the quota was based on \hat{B}_y , because the fishery selects some immature ages. The fully selected fishing mortality rate that would remove the quota from the true population (\tilde{F}_y) was found using Newton–Raphson iterations. Several variations of the biomass-based rule were also evaluated. These variations included applying the control rule annually, using the same quota for 3-year blocks such that the control rule is applied every fourth year (i.e., $Q_{y+1} = Q_{y+2} = Q_{y+3}$), using the same quota for 5-year blocks, and using the same quota for 3-year blocks but restricting the change in the quota to 15% in either direction when the control rule was reapplied in the fourth year. Thus, four variants of the biomass based control rule were evaluated: (i) annual application, (ii) 3-year blocks, (iii) 5-year blocks, and (iv) 3-year blocks with a 15% restriction.

For each biomass-based control rule variant, a range of values for the three parameters defining the control rule was evaluated. The proportion (ψ) of F_{MSY} that dictates the maximum desired fishing mortality rate was varied from $0.1F_{MSY}$ to $1.0F_{MSY}$ in increments of 0.1, while the lower and upper SSB threshold parameters (SSB_{low} and SSB_{up}) were varied from $0.0SSB_{MSY}$ to $4.0SSB_{MSY}$ but with inconsistent increments (i.e., 0.0, 0.1, 0.3, 0.5, 0.7, 0.9, 1.0, 1.1, 1.3, 1.5, 1.7, 2.0, 2.5, 3.0, 3.5, 4.0). The full factorial of combinations for the three biomass-based control rule parameters produced 1360 shapes (note SSB_{low} must be $\leq SSB_{up}$), and each of these shapes was evaluated for each of the four biomass-based control rule variants described above.

The constant catch control rule is defined by one parameter, a desired constant catch (i.e., quota) amount. The constant catch amounts were varied from $0.1MSY$ to $1.0MSY$ in increments of 0.1.

The conditional constant catch rule used a constant desired catch amount unless removing that desired catch from the assessed biomass caused the fully selected fishing mortality rate to exceed a predetermined maximum, in which case the desired catch was set to the value produced by applying the maximum fully selected fishing mortality rate to the assessed biomass. Thus, the conditional constant catch rule has two policy parameters: a desired constant catch amount and a maximum fishing mortality rate. The constant catch amounts were varied over the same range as in the constant catch control rule, while the maximum fishing mortality rate equaled $0.5F_{MSY}$. When the maximum fishing mortality rate portion of the conditional constant catch rule was invoked, a quota was set in the same manner as when $\overline{SSB}_y \geq SSB_{up}$ in the biomass-based control rule described above.

Implementation error

Implementation errors were also included in a similar way as in Punt et al. (2008) and Deroba and Bence (2012), as year-specific lognormal random deviations:

$$(9) \quad F_{a,y} = \bar{F}_y S_a e^{\frac{\varepsilon_{\theta,y} - \sigma_{\theta}^2}{2}} \quad \varepsilon_{\theta} \sim N(0, \sigma_{\theta}^2)$$

The variance of implementation errors (σ_{θ}^2) equaled 0.001. The US Atlantic herring fishery in the Northwest Atlantic generally catches the full amount of annual quota. Catches are monitored

through mandatory federal and state reporting requirements, which are used to close the fishery within 10% of the annual quota. Thus, unbiased implementation errors seemed justified.

Predators

The food web of the Northeast US continental shelf large marine ecosystem is characterized by many diverse predators and prey (Link 2002). There is a wealth of scientific information to characterize predator–prey relationships in this region, including feeding ecology data for fish predators (e.g., Smith and Link 2010), seabirds (Hall et al. 2000), bluefin tuna (*Thunnus thynnus*) (Chase 2002; Golet et al. 2013, 2015; Logan et al. 2015), and marine mammals (Smith et al. 2015). Consumption of herring by predators has been extensively studied in this ecosystem (Overholtz et al. 2000; Overholtz and Link 2007), and multiple methods were evaluated to include this consumption within the most recent herring benchmark stock assessment (Northeast Fisheries Science Center 2012).

Much of this information was presented at the first stakeholder workshop in May 2016, where it was agreed that separate “general predator” models linked to herring would be a reasonable approach, with the goal of developing one model for each of the four predator categories: highly migratory fish, groundfish, seabirds, and marine mammals (Feeney et al. 2019). Bluefin tuna were identified at the stakeholder workshop as a recommended highly migratory herring predator, and common terns (*Sterna hirundo*) were identified at the stakeholder workshop as the recommended seabird herring predator, so these predators were modeled. No specific groundfish or marine mammal was identified as a representative herring predator during the stakeholder workshop. In sections below, we discuss these decisions further.

Predators were modeled with fairly simple delay difference population dynamics that allowed different predator population processes to be dependent on some aspect of herring population status, following Plagányi and Butterworth (2012). Each predator model takes output from the herring operating model as input, and outputs performance metrics identified at the stakeholder workshop. While this allows “bottom-up” effects of herring on predators to be examined, this configuration does not consider “top-down” effects of predators on herring or simultaneous interactions of multiple predators with herring.

There were two modeling components for each predator included in the herring MSE: a predator population model and a herring–predator relationship model to link herring with predator populations. Here, we give an overview of the modeling process, and we describe the decisions made in parameterizing individual predator models and herring–predator relationships in the following sections. The overall population in numbers for each predator P each year N_y^P is modeled with a delay difference function:

$$(10) \quad N_{y+1}^P = N_y^P S_y^P + R_{y+1}^P$$

where annual predator survival S_y^P is based on annual natural mortality v and exploitation u

$$(11) \quad S_y^P = (1 - v_y)(1 - u)$$

and annual recruitment R_y^P (delayed until recruitment age a) is a Beverton–Holt function defined as above for herring.

Predator population biomass is defined with Ford–Walford plot intercept (FW_{int}) and slope (FW_{slope}) growth parameters

$$(12) \quad B_{y+1}^P = S_y^P (FW_{int} N_y^P + FW_{slope} B_y^P) + FW_{int} R_{y+1}^P$$

Parameterizing this model requires specification of the stock–recruitment relationship (steepness h and unfished spawning stock size in numbers or biomass), the natural mortality rate, the fishing mortality (exploitation) rate, the initial population size,

Table 5. Predator population model specification and parameter sources.

	Highly migratory	Seabird	Groundfish	Marine mammal
Stakeholder preferred species	Bluefin tuna	Common tern	Not specified	Not specified
Species modeled	Bluefin tuna (western Atlantic stock)	Common tern (GOM colonies as defined by the GOM Seabird Working Group)	Spiny dogfish (GOM and GB cod stocks also examined)	None; data limited (minke and humpback whales, harbor porpoise, harbor seal examined)
Stock–recruitment	ICCAT 2015 and Porch and Laretta 2016	Derived from GOMSWG data	Rago et al. 1998 and Rago and Sosebee 2010	No time series data for our region
Natural mortality	ICCAT 2015	Nisbet 2002	Rago and Sosebee 2015 and P. Rago, personal communication, 2016	Derivable from Hayes et al. 2017?
Fishing mortality	ICCAT 2015	NA	Rago and Sosebee 2015 and P. Rago, personal communication, 2016	Derivable from Hayes et al. 2017?
Initial population	ICCAT 2015	Derived from GOMSWG data	Rago and Sosebee 2015 and P. Rago, personal communication, 2016	Derivable from Hayes et al. 2017?
Weight-at-age	Restrepo et al. 2010	Nisbet 2002	Rago et al. 1998	Depending on species, derivable from literature

Note: See methods for details of Gulf of Maine Seabird Working Group (GOMSWG) data. GOM, Gulf of Maine; GB, Georges Bank.

and the weight-at-age of the predator (Ford–Walford plot intercept and slope parameters). For each predator, population parameters were derived from different sources (Table 5).

Predator population models were based on either the most recent stock assessment for the predator or from observational data from the Northeast US shelf. Herring–predator relationships were based on either peer-reviewed literature or observational data specific to the Northeast US shelf. We did not include process or observation error in any of these modeled relationships. This is obviously unrealistic, but the primary objective of the herring MSE is to evaluate the effect of herring management on predators. Leaving out variability driven by anything other than herring is intended to clarify the effect of herring management.

To develop the herring–predator relationship model, specific herring population characteristics (e.g., total abundance or biomass or abundance–biomass of certain ages or sizes) were related to predator growth, predator reproduction, or predator survival. Our aim was to use information specific to the Northeast US shelf ecosystem from peer-reviewed literature, from observations, or a combination.

In general, if support for a relationship between herring and predator recruitment was evident, it was modeled as a predator recruitment multiplier based on the herring population (N_y) relative to a specified threshold N_{thresh} :

$$(13) \quad \bar{R}_{y+a}^p = R_{y+a}^p \cdot \frac{\gamma(N_y/N_{\text{thresh}})}{(\gamma - 1) + (N_y/N_{\text{thresh}})}$$

where $\gamma > 1$ links herring population size relative to the threshold level to predator recruitment.

If a relationship between predator growth and herring population size was evident, annual changes in growth were modeled by modifying either the Ford–Walford intercept (α_y^p) or slope (ρ_y^p):

$$(14) \quad B_{y+1}^p = S_y^p(\alpha_y^p N_y^p + \text{FW}_{\text{slope}} B_y^p) + \alpha_y^p R_{y+1}^p$$

or

$$(15) \quad B_{y+1}^p = S_y^p(\text{FW}_{\text{int}} N_y^p + \rho_y^p B_y^p) + \text{FW}_{\text{int}} R_{y+1}^p$$

where either α_y^p or ρ_y^p are defined for a predator using herring population parameters (see eq. 17 below).

Finally, herring population size N_y could be related to predator survival using an annual multiplier on constant predator annual natural mortality v :

$$(16) \quad v_y = v e^{-\left(\frac{N_y}{N_{F=0}}\right)\delta}$$

where $0 < \delta < 1$ links herring population size to predator survival.

After specifying the population model parameters and herring–predator relationship, we applied the (Hilborn and Walters 2003) equilibrium calculation for the delay difference model with $F = 0$ to get the unfished spawners per recruit ratio. This ratio was then used in a second equilibrium calculation with the current predator exploitation rate to estimate Beverton–Holt stock–recruitment parameters, equilibrium recruitment, and equilibrium individual weight under exploitation. Then, each model was run forward for 150 years with output from the herring operating model specifying the herring population characteristics.

Tuna population model

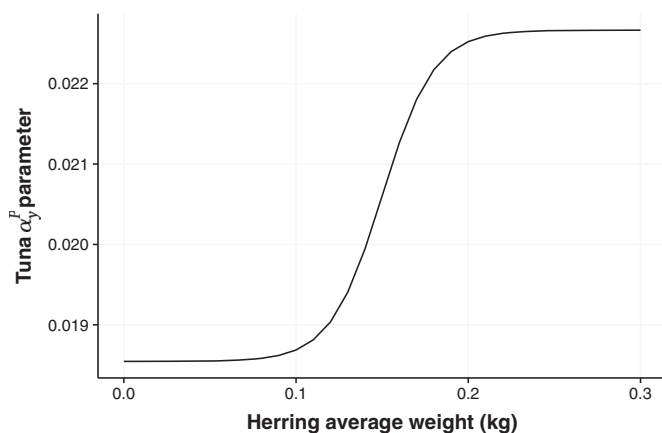
Western Atlantic bluefin tuna population parameters were drawn from the 2014 stock assessment (ICCAT 2015), the growth curve from Restrepo et al. (2010), and recruitment parameters from a detailed examination of alternative stock–recruit relationships (Porch and Laretta 2016). Ultimately, the “low recruitment” scenario was selected to represent bluefin tuna productivity in the Gulf of Maine, which defines B_{MSY} as 13 226 t and therefore affects measures of status relative to B_{MSY} . Continuation of the current tuna fishing strategy ($F < 0.5F_{\text{MSY}}$ under the low recruitment scenario) is assumed. All predator population model parameters are listed in Table 6.

Herring–tuna relationship model

Tuna diets are variable depending on location and timing of foraging (Chase 2002; Golet et al. 2013, 2015; Logan et al. 2015), but for the purposes of this analysis, we assumed that herring is an important enough prey of tuna to impact tuna growth in the Northeast US shelf ecosystem. A relationship between bluefin tuna growth and herring average weight was implemented based on information and methods in Golet et al. (2015). The relationship between tuna condition anomaly (defined as proportional departures from the weight-at-length relationship used in the assessment) and average weight of tuna-prey-sized herring (\bar{W}_y , herring >180 mm collected from commercial herring fisheries)

Table 6. Predator model input parameters.

Parameter	Tuna	Tern	Dogfish
Numbers or weight	Weight	Numbers	Weight
Unfished spawning population (metric tons for fish, n nesting pairs for tern)	66 900	45 000	300 000
Steepness h	1.0	0.26	0.97
Base annual natural mortality rate ν	0.14	0.1	0.092
Annual exploitation rate u	0.079	0.0	0.092
Growth intercept FW_{int}	0.020605	0.00015	0.000278
Growth slope FW_{slope}	0.9675	0.0	0.9577
Initial abundance N_1^p (n for fish, n nesting pairs for tern)	111 864	3 000	49 629 630
Initial biomass B_1^p (metric tons for fish, kilograms for tern but not used in model)	27 966	1.5	134 000
Recruit delay (age, years)	1	4	10
Prey-recruitment link γ	1 (off)	1.09	1 (off)
Prey-mortality link δ	0 (off)	0 (off)	0.2
Prey-growth link λ	1.1	1 (off)	1 (off)

Fig. 1. Modeled herring average weight (population > 180 mm in length) to tuna growth relationship. See text (Herring–tuna relationship model section) for derivation.

was modeled as a generalized logistic function with lower and upper bounds on tuna growth parameters:

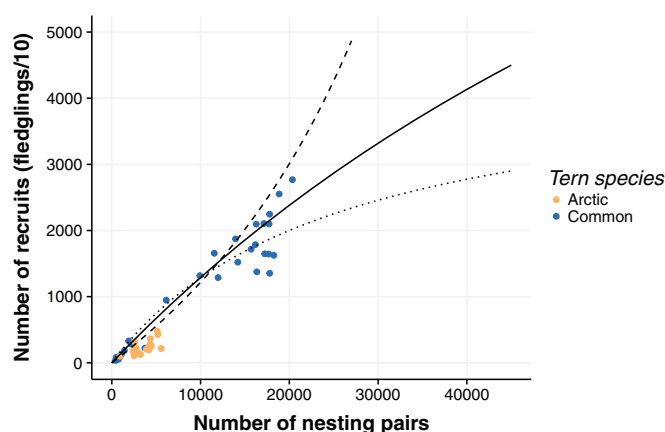
$$(17) \quad \alpha_y^p = (0.9FW_{int}) + \frac{(1.1FW_{int}) - (0.9FW_{int})}{1 + e^{(1-\lambda) \cdot [100(\bar{W}_y - T)/T]}}$$

where $\lambda > 1$ links herring average weight anomalies to tuna growth.

The inflection point of $T = 0.15$ kg average weight aligns with 0 tuna weight anomaly from figure 2C on p. 186 in Golet et al. (2015), and upper and lower bounds were determined by estimating the growth intercept with weight-at-age 10% higher or lower, respectively, from the average weight-at-age obtained by applying the length to weight conversion reported in the 2014 stock assessment (ICCAT 2015) to the length-at-age estimated from the Restrepo et al. (2010) growth curve (Fig. 1). When included in the model with $\lambda = 1.1$ in eq. 17, the simulated variation in tuna weight-at-age covered the observed range reported in Golet et al. (2015).

Tern population model

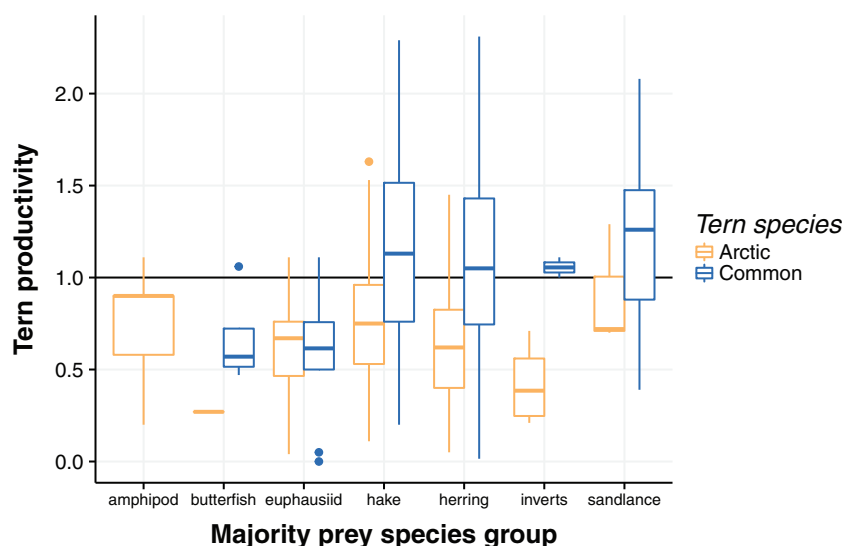
There is no published stock assessment or population model for most seabirds in the Northeast US. Therefore, Gulf of Maine common and Arctic tern (*Sterna paradisaea*) population parameters were drawn from accounts in the Birds of North America web page (Hatch 2002; Nisbet 2002) and estimated from counts of breeding pairs and estimates of fledgling success summarized by the Gulf of Maine Seabird Working Group (GOMSWG; data at <http://gomswg.org/minutes.html>), as corrected and updated by seabird experts from throughout Maine. While we initially ana-

Fig. 2. Stock–recruitment function for Gulf of Maine common terns assuming 10% fledgling to adult survival. Fitted parameters with all years of the common tern data set included a nonsignificant beta parameter (dashed line), while fits to a truncated data set resulted in low population production rates inconsistent with currently observed common tern trends (dotted line). Therefore, steepness was estimated to give a relationship (solid black line) falling between these two lines. [Colour online.]

lyzed both Arctic and common tern information, the stakeholder workshop identified common terns as the example species for modeling, and this species has more extensive data and a generally higher proportion of herring in its diet based on that data. Therefore, this predator model is based on common terns in the Gulf of Maine.

Adult breeding pairs by colony were combined with estimated productivity of fledglings per nest to estimate the annual number of fledglings for each year. A survival rate of 10% was applied to fledglings from each year to represent “recruits” to the breeding adult population age-4 and up (Nisbet 2002). This “stock–recruit” information was used to estimate steepness for the delay difference model based on common tern information only. Fitting parameters with R nls (R Core Team 2016) had variable success, with the full data set unable to estimate a significant beta parameter (dashed line, Fig. 2) for common terns and a truncated data set resulting in low population production rates inconsistent with currently observed common tern trends (dotted line, Fig. 2). Therefore, steepness was estimated to give a relationship (solid line, Fig. 2) falling between these two lines. The resulting stock–recruit relationship set steepness at 0.26, a theoretical maximum breeding adult population of 45 000 pairs (1930s New England population; Nisbet 2002) and a theoretical maximum recruitment of 4500 individuals annually (reflecting approximately a productivity of 1.0 at “carrying capacity” resulting in a stable population).

Fig. 3. Gulf of Maine tern annual productivity distributions by majority diet item offered to fledglings. Boxplots represent the median (wide line within the box) and 25th and 75th percentiles (box) of annual productivity measured across all nesting colonies in the Gulf of Maine where that prey species was the majority in the diet. Boxplot whiskers include the highest and lowest observations within 1.5 box lengths from the box. Observations further from the box (outliers) are represented by points. The horizontal black line represents the target tern productivity of 1.0 fledgling per nest. [Colour online.]



Average common tern productivity is 1.02 (all Gulf of Maine colony data combined). Adult mortality was assumed to be 0.1 for the delay difference model (survival of 90% for adults; Nisbet 2002). The resulting model based on common tern population dynamics in the Gulf of Maine (with no link to herring) predicts that the population will increase to its carrying capacity under steady conditions over a 150-year simulation. The actual population has increased at ~2% per year between 1998 and 2015 (GOMSWG data). Given the lack of detailed demographic information in the delay difference model, this was considered a good representation of the average observed trend in current common tern population dynamics.

Herring–tern relationship model

The relationship between herring abundance and tern reproductive success was built based on information from individual colonies on annual productivity, proportion of herring in the diet, and amount of herring in the population as estimated by the current stock assessment. Since little of this information has appeared in the peer-reviewed literature, we present it in detail here. First, productivity information was evaluated by major diet item recorded for chicks over all colonies and years. In general, common tern productivity was higher when a streamlined fish species (“hake”, “herring”, and “sandlance” in Fig. 3) was the major diet item relative to invertebrates (“amphipod”, “euphausiid”, and “inverts” = unidentified invertebrates in Fig. 3). However, having herring as the major diet item resulted in about the same distribution of annual productivities as having unidentified juvenile hake (*Urophycis* or *Merluccius* spp.) or sandlance (*Ammodytes* spp.) as the major diet item for these colonies (Fig. 3).

Individual colonies showed different trends in number of nesting pairs, productivity, and proportion of herring in the diet (plots available upon request). When both Arctic and common terns shared a colony, interannual changes in productivity were generally similar between species, suggesting that conditions at and around the colony (weather, predation pressure, and prey fields) strongly influenced productivity rather than species-specific traits. Only two colonies (Machias Seal Island near the US–Canada border and Stratton Island in southern Maine) showed a significant positive correlation between the proportion of herring in the chick diet and productivity (Machias Seal: Spearman’s rank rho =

0.63, $p = 0.019$; Stratton: Spearman’s rank rho = 0.52, $p = 0.035$). Other islands showed either nonsignificant (no) relationships or in one case (Metinic Island) a significant negative relationship (Fig. 4).

The estimated population size of herring on the Northeast US shelf had some relationship to the amount of herring in tern diet at several colonies (4 of 13 common tern colony diets related to herring age-1 recruitment, 6 of 13 common tern colony diets related to herring total biomass (B), and 4 of 13 common tern colony diets related to herring SSB; detailed statistics and plots available upon request). However, statistically significant direct relationships between herring population size and tern productivity were rare, with only Ship Island productivity increasing with herring total B , and Eastern Egg Rock, Matinicus Rock, Ship, and Monomoy islands productivity increasing with herring SSB. Given that Monomoy Island tern chicks consistently displayed the lowest proportion of herring in their diets of any colony (0%–11%), we did not consider this relationship further to build the model.

Based on tern feeding observations, we would expect the number of age-1 herring in the population to be most related to tern productivity since that is the size class terns target, but this relationship was not found in analyzing the data. Herring total biomass was positively related to tern diets at nearly half of the colonies and reflects all size classes including the smaller sizes most useful as tern forage, but was only directly related to tern productivity at one colony. Herring SSB was not considered further as an index of tern prey because it represents sizes larger than tern forage.

To represent the potential for herring to influence tern productivity, we parameterized a tern “recruitment multiplier” based on herring assessed total biomass and common tern productivity across all colonies (except Monomoy Island where terns eat sandlance). This relationship includes a threshold herring biomass where common tern productivity would drop below 1.0, and above that threshold productivity exceeds 1.0 (Fig. 5). The threshold of ~400 000 t is set where a linear relationship between herring total biomass and common tern productivity crosses productivity = 1 (black dashed line in Fig. 5). However, the selected threshold is uncertain because there are few observations of common tern productivity at low herring total biomass (1975–1985).

Fig. 4. Herring proportion in diet and annual tern productivity (fledglings/nest) by Gulf of Maine colony. Colony names are at the top of each subplot; I = island. Two colonies, Machias Seal and Stratton, have significant positive Spearman's rank correlations between herring proportion in diet and annual productivity for common terns. No linear model slopes were significant, so none are shown in the plot. The horizontal black line represents the target tern productivity of 1.0 fledgling per nest. [Colour online.]

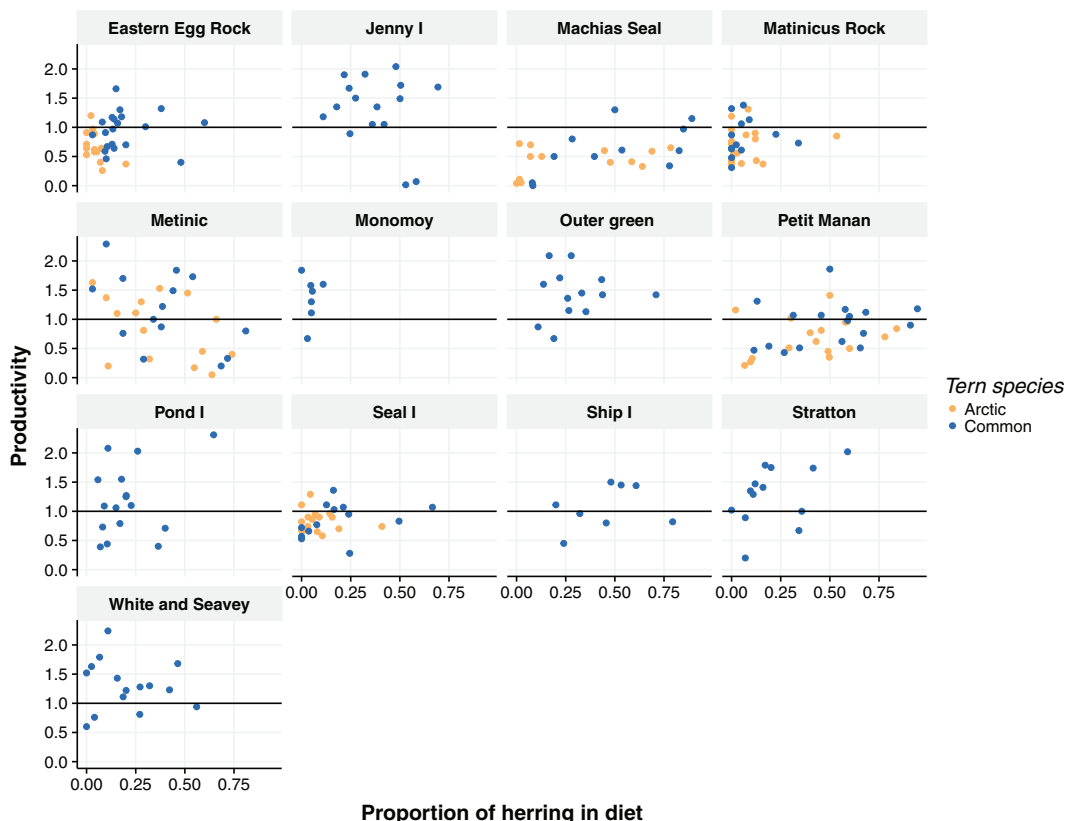
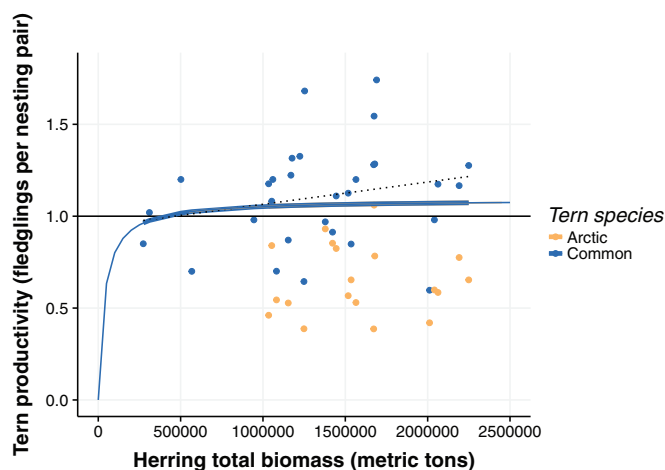


Fig. 5. Modeled influence of herring total biomass on tern reproductive success. Total annual productivity (fledglings/nest) for both tern species relative to assessed herring total biomass is shown, but the modeled relationship (curve) is based only on common terns. A linear relationship between herring total biomass and common tern productivity crosses tern productivity = 1 (black dotted line) at 400 000 t herring total biomass. This linear relationship does not have a statistically significant slope; the curve was fit to represent a level but positive contribution of herring total biomass to common tern productivity above the threshold. The horizontal black line represents the target tern productivity of 1.0 fledgling per nest. [Colour online.]



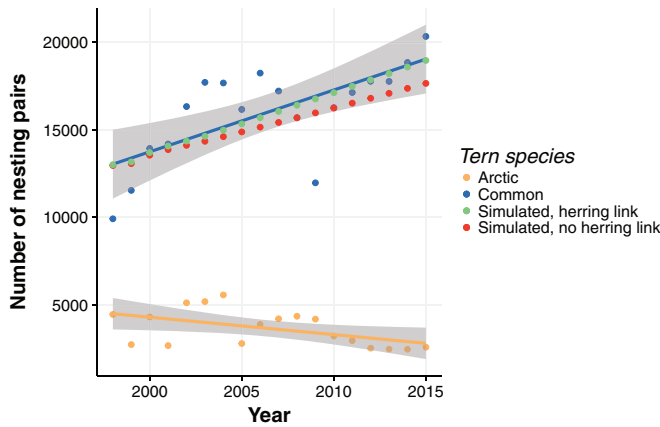
The linear relationship does not have a statistically significant slope; a curve was fit to represent a level contribution of herring total biomass to common tern productivity above the threshold. The curve descends below the threshold, dropping below 0.5 productivity at around 50 000 t and representing the extreme assumption that herring extinction would result in tern productivity of 0. Although the relationship of tern productivity to herring biomass at extremely low herring populations has not been quantified, control rules that allow herring extinction do not meet stated management objectives for herring, so this extreme assumption for terns will not change any decisions to include or exclude control rules.

When included in the model using $\gamma = 1.09$ in eq. 13, this relationship adjusts the modeled common tern population increase to match the current average increase in common tern nesting pairs observed in the data (Fig. 6). There is still considerable uncertainty around this mean population trajectory, which cannot be reflected in our simple model.

Groundfish

Because no specific groundfish was identified as a representative herring predator during the stakeholder workshop, the first decision was which groundfish to model. Annual diet estimates (based on sample sizes of ~100+ stomachs) are available for the top three groundfish predators of herring (those with herring occurring in the diets most often in the entire NEFSC food habits database): spiny dogfish (*Squalus acanthias*, hereinafter dogfish), Atlantic cod (*Gadus morhua*, hereinafter cod), and silver hake (*Merluccius bilinearis*). Cod and spiny dogfish were considered first because their overall diet proportions of herring are higher and because silver hake has the least recently updated assessment.

Fig. 6. Population trends for Gulf of Maine terns with and without the simulated herring–common tern productivity relationship. Linear model fit (line) with 95% CI band (gray shading) is shown for significant relationships. [Colour online.]



Diet compositions by year were estimated for spiny dogfish, Georges Bank cod, and Gulf of Maine cod to match the scale of stock assessments. Full weighted diet compositions were estimated and suggest considerable interannual variability in the herring proportion in groundfish diets (filled proportions of bars in Fig. 7).

Some interannual variation in diet may be explained by changing herring abundance. Dogfish and both cod stocks had positive relationships between the amount of herring observed in annual diets and the size of the herring population according to the most recent assessment (statistics and plots available upon request). This suggests that these groundfish predators are opportunistic, eating herring in proportion to their availability in the ecosystem. However, monotonically declining cod populations for both Gulf of Maine and Georges Bank cod stocks resulted in either no herring–cod relationship or a negative relationship between herring populations and cod populations (Fig. 8). Only dogfish spawning stock biomass had a positive relationship with the proportion of herring in dogfish diet. Therefore, we selected dogfish as the groundfish predator for modeling.

Dogfish population model

The dogfish model stock–recruitment function, initial population, and annual natural mortality were adapted from information in Rago et al. (1998), Rago and Sosebee (2010), Bublely et al. (2012), and Rago and Sosebee (2015). Because of differential growth and fishing mortality by sex, our model best represents female dogfish (a split-sex delay difference model was not feasible within the time constraints of this MSE). Further, dogfish stock–recruit modeling to date based on Ricker functions (Rago and Sosebee 2010) captures more nuances in productivity than the Beverton–Holt model we used. Our recruitment parameterization reflects a stock with generally low productivity and relatively high resilience, which we recognize is a rough approximation for a species such as dogfish. The annual fishing exploitation rate applied is the average of the catch/adult female biomass from the most recent years of the 2016 data update provided to the Mid-Atlantic Fishery Management Council (P. Rago, personal communication, 2016).

Herring–dogfish relationship model

There was a weak positive relationship between dogfish total biomass and herring total biomass from the respective stock assessments (Spearman’s rank correlation = 0.36, $p = 0.012$; Pearson’s correlation = 0.32, $p = 0.026$), but no significant relationship between dogfish weight or dogfish recruitment and herring population size. During the recent period of relatively low dogfish

recruitment (1995–2007), there was a positive relationship between juvenile dogfish (pup) average weight and herring proportion in diet, suggesting a potential growth and (or) recruitment mechanism; however, this relationship does not hold throughout the time series, so we considered it too weak as a basis for population modeling.

Therefore, to simulate a potential positive relationship between herring and dogfish, we assumed that dogfish survival increased (natural mortality was reduced) by an unspecified mechanism as herring abundance increased (Fig. 9). Because dogfish are fully exploited by fisheries in this model, the impact of this change in natural mortality on total survival has small to moderate benefits to dogfish population numbers and biomass. Using a $\delta = 0.2$ in eq. 16 results in weak increases in dogfish biomass with herring abundance consistent with observations.

Marine mammals

Because no specific marine mammal was identified as a representative herring predator in the stakeholder workshop, as with groundfish, the first decision was which marine mammal to model. Diet information for a wide range of marine mammals on the Northeast US shelf suggests that minke whales (*Balaenoptera acutorostrata*), humpback whales (*Megaptera novaeangliae*), harbor seals (*Phoca vitulina*), and harbor porpoises (*Phocoena phocoena*) have the highest proportions of herring in their diets (Smith et al. 2015) and therefore may show some reaction to changes in the herring HCR.

While some food habits data existed for marine mammals, consultation with marine mammal stock assessment scientists at the Northeast Fisheries Science Center confirmed that no data were available to parameterize a stock–recruitment relationship for any of these marine mammal species in the Northeast US region, and no such information was available in the literature for stocks in this region. Although it may be possible to develop stock–recruitment models for one or more of these species in the future, it was not possible within the time frame of the herring MSE. Therefore, we were unable to model marine mammals within the same framework as other predators.

Potential effects of changes in herring production and (or) biomass on marine mammals were instead evaluated using an updated version of an existing food web model for the Gulf of Maine (Link et al. 2006, 2008, 2009) and incorporating food web model parameter uncertainty. Overall, food web modeling showed that a simulated increase in herring production in the Gulf of Maine might produce modest but uncertain benefits to marine mammal predators, primarily because increased herring was associated with decreases in other forage groups also preyed on by marine mammals. However, this could not be pursued further within the MSE framework.

Predator model input parameters are summarized in Table 6.

Economics

The herring fishery

The economic model of the herring fishery converts yield, Y , from the herring model component into gross revenues (GR) and net operating revenues (NR). There are two fleets, trawl and purse seine, that are assumed to have the ability to catch 70% and 30% of the yield, respectively. This division corresponds to recent historical patterns. The midwater trawl, paired midwater trawl, and bottom trawl are all aggregated into the trawl fleet. GR, NR, and the constraints on harvest can be represented as follows (year subscripts omitted here for simplicity):

$$(18) \quad GR = p(q^t + q^s)q^t + p(q^t + q^s)q^s$$

$$(19) \quad NR = GR - c^t(q^t) - c^s(q^s)$$

Fig. 7. Annual percentage of herring in diet compositions for major groundfish predators of herring (dogfish and two cod stocks) estimated from NEFSC food habits database. Prey identified to herring family (Clupeidae, open bars) as well as Atlantic herring (filled bars) are included. [Colour online.]

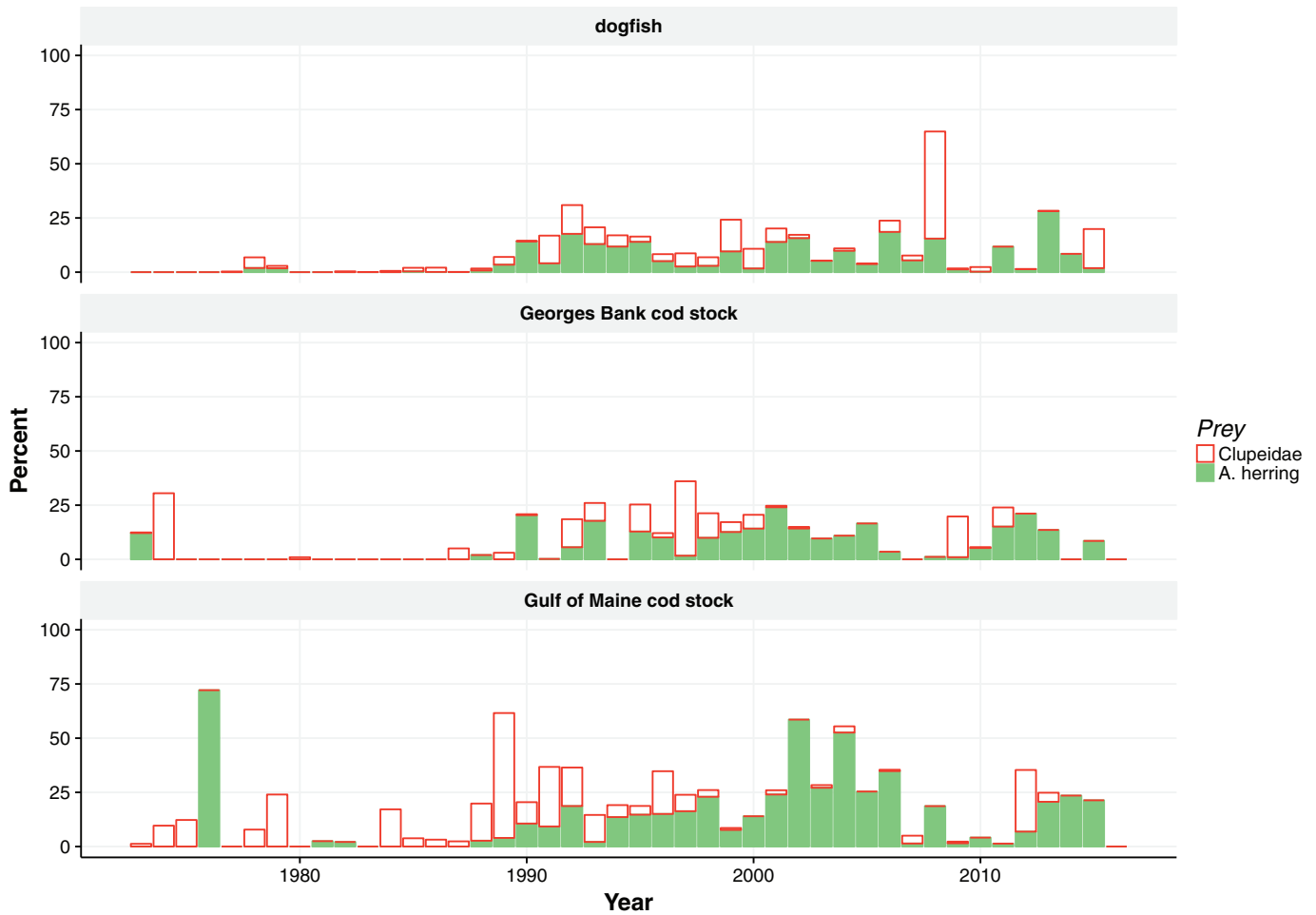


Fig. 8. Relationship of groundfish predator (dogfish, GBcod = Georges Bank cod, and GOMcod = Gulf of Maine cod) spawning stock biomass (SSB) with the percentage of herring in diet: all clupeids including Atlantic herring (All Clupeids) and Atlantic herring only (A. herring). Linear model fit (line) with 95% CI band (gray shading) is shown for significant relationships. [Colour online.]

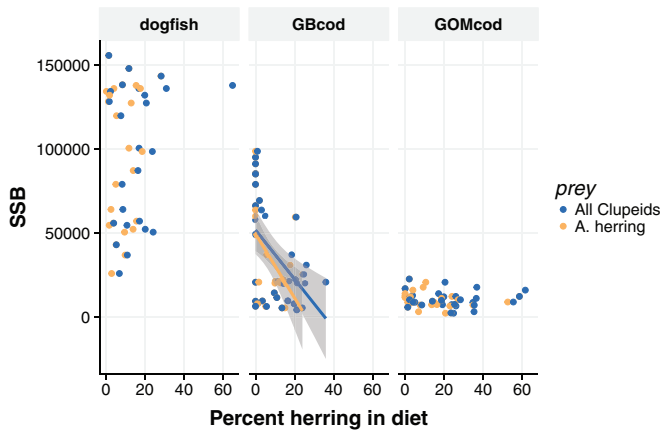
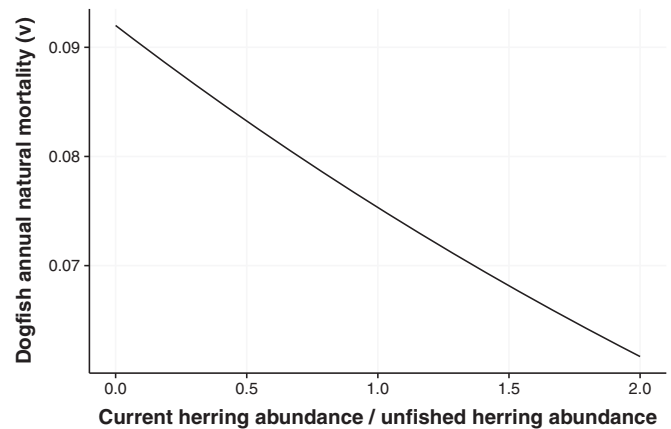


Fig. 9. Modeled herring relative population size – dogfish natural mortality relationship.



where q^i is the quantity landed for fleet i , $c^i(q^i)$ is cost function for fleet i , and $p(\cdot)$ is a function that relates total landings to prices.

$$(20) \quad \max_{q_i} NR^i = p(q^t + q^s)q^i - c^i(q^i)$$

$$(21) \quad q^s \leq 0.3\text{yield}; \quad q^t \leq 0.7\text{yield}$$

Table 7. Marginal cost of 1 t of herring, daily catch, trip length, and adjusted cost per day for the purse seine and trawl fleets averaged over 2011–2015.

Year	Marginal cost (\$·t ⁻¹)	Catch (t·day ⁻¹)	Trip length (days)	Adjusted cost (\$·day ⁻¹)	Observed trips	VTR trips
Seine	14.27	87.5	1.0	1249	207	1413
Trawl	62.43	61.4	2.9	3833	573	2005

Note: VTR, vessel trip report.

The optimization problem in eq. 20 contains two embedded assumptions: total catch is less than or equal to yield, and a fleet may catch less than its fraction of yield (presumably because it may be more profitable to select a lower level of landings).

Economic data collected from 2011 to 2015 by the Northeast Fisheries Observer Program were used to construct average daily costs for the trawl and purse seine fleets. Fuel prices were much lower in 2011–2014 compared with 2015. We adjusted fuel prices to the 2011–2014 average; sensitivity analysis was performed by setting fuel prices to the 2015 levels, but results are not reported here. Other costs of fishing included water, oil, and damage costs. Crew pay and fixed costs were not included.

We construct average catch per day fished for each fleet from the vessel trip report databases over the same time period. The trip lengths in the vessel trip report and observer data were similar. This allows us to construct the average cost of catching a metric ton of herring for the trawl and purse seine fleets (c^t and c^s , respectively). We assume that the average cost is equal to the marginal cost for each fleet. These figures are presented in Table 7.

Annual prices were constructed from NMFS dealer data for 1982 through 2016. Annual landings were constructed from the processed Maine Department of Marine Resources landings data set for the same time period. Prices have been normalized to 2015 real US dollars using the Bureau of Labor Statistics Producer Price Index for Unprocessed and Packaged Fish (WPU0223). Because Atlantic herring was not federally managed prior to the implementation of the Herring Fisheries Management Plan in 2000; the NMFS dealer databases may not contain all landings prior to this time. The Maine Department of Marine Resources data do not contain prices but is a census of landings. Exploratory analysis suggested both a regime change in the mid-1990s and likely non-stationarity of both landings and prices. We used the testing methodology developed by Pesaran et al. (2001) to examine the existence of a long-run relationship between prices and quantities. This method does not require pretesting for stationarity; however, the test statistic does have an inconclusive zone in which knowledge of stationarity would be required. A long-run relationship between prices and landings can be modeled as follows:

$$(22) \quad p_y = c + \sum_{i=1}^p a_i p_{t-i} + \sum_{i=0}^n b_i q_{y-i} + e_t$$

or equivalently as an error correction model:

$$(23) \quad \Delta p_y = \gamma + \alpha_1 p_{y-1} + \sum_{i=1}^p \theta_i \Delta p_{y-i} + \sum_{i=0}^n \beta_i \Delta q_{y-i} + e_y$$

where Δ is the first-differences operator (Pesaran et al. 2001). The γ parameter must also be restricted ($\gamma = c/\alpha_1$) for eqs. 22 and 23 to be equivalent. Pesaran et al. (2001) tests the null of no long-run relationship using a joint F test evaluating whether the α_i and β_i parameters in eq. 23 are nonzero; however, the F statistic has a nonstandard distribution with an inconclusive area.

Equation 23 was first estimated on the full 1982–2015 data set; model selection criteria indicated a model with four lags of price and no lags of quantities ($p = 4, n = 0$) was preferred and that prices

Table 8. Regression results from the autoregressive distributed lag (1,0) specification (eq. 22), Pesaran–Shin–Smith (PSS; Pesaran et al. 2001) bounds test and associated critical values.

	Model	
	(1) price	(2) ln price
Price _{t-1}	0.646*** (-0.0937)	0.666*** (-0.0935)
Quantity	-1.194*** (-0.277)	-0.395*** (-0.0956)
Constant	217.8*** (-48.16)	6.423*** (-1.484)
Observations	21	21
R ²	0.906	0.906
BGp	0.34	0.547
BGF	1	0.4
PSS F statistic	10.2	9.34
1% critical values	(6.84, 7.84)	

Note: Standard errors are given in parentheses. Data are for 1995–2015 only. The explanatory variables enter in levels in the first column (1) and natural logarithms in the second column (2). ***, $p < 0.01$.

and quantities had no long-run relationship. We suspect this is likely caused by a combination of overfitting of the model and a regime shift evident in the exploratory graphs. Rather than explore a regime switching model, we simply estimated eq. 23 with $p = 1, n = 0$ on a subset of the data (1995–2015). If there was a regime shift, the current regime is more likely to be similar to the future. Models estimated in natural logarithms and in levels fit well. The Pesaran–Shin–Smith (Pesaran et al. 2001) F statistics of 9.34 and 10.20 are above the upper critical value, strongly suggesting a long-run relationship between prices and quantities (Table 8). We also present the results of the autoregressive distributed lag (1,0) formulation because it is a bit easier to interpret. As a robustness check, we also tried varying the first (1996) and last (2016) year included in the data set. This did not change the estimated results substantially. We also estimated a short-run relationship between prices and quantities in which Δp_y was regressed on Δq_y . The short-run effects were qualitatively similar to the long-run model in Table 8. Coefficients from the “level” equation (column (1) of Table 8) are used in the simulation.

The simulation of gross operating revenues occurs in a few steps. Herring prices in a year are simulated using eqs. 20 and 22 and parameters from Table 8 previous year prices (initialized to the 2011–2015 average for the first year) under the assumption that both fleets combine to land the entire yield. Following Lehuta et al. (2014), if the price of herring is sufficiently high, we assume that consumers find it worthwhile to switch to Atlantic menhaden (*Brevoortia tyrannus*). If the simulated prices are higher than the price of menhaden (\$242/t) plus transport costs (\$133/t), we set the price of herring to \$375/t. When simulated prices are higher than the marginal cost of the trawl fleet (\$63.24/t), both fleets are assumed to catch the entire yield. Otherwise, we use eqs. 20 and 22 to solve for quantity landed by the trawl fleet when the purse seine fleet lands 30% of yield. If it is optimal for the trawl fishery to land nothing, we use eqs. 20–22 to find the purse seine’s optimal amount of landings. Because the marginal costs for the purse seine are always less than the marginal costs of the trawl fishery, any landings by the trawl fleet imply the purse seine fleet is landing 30% of the yield. NR to the fishery can then be calculated directly from eq. 20.

Performance metrics

For each combination of control rule and operating model (43 680 unique combinations), 100 simulations lasting 150 years were conducted. Preliminary simulations suggested that this

number of simulations and years was sufficient for results to be insensitive to starting conditions and short-term dynamics caused by autocorrelated processes. The simulated herring time series for every operating model and control rule was passed to the predator and economic submodels, resulting in outputs as described below using the equations above. We report performance metrics over the final 50 years as a way to describe the long-run performance of a particular control rule. Performance metrics were derived directly from the results of the stakeholder workshop and supplemented with additional metrics drawn from MSE best practices (Punt et al. 2016a).

Herring performance metrics

Median SSB, $\frac{SSB}{SSB_{F=0}}$, $\frac{SSB}{SSB_{MSY}}$, yield, $\frac{yield}{MSY}$, biomass of herring dying due to M , and the proportion of the herring population composed of age-1 fish were recorded as herring performance metrics. Additional performance metrics included the proportion of years with $SSB < SSB_{MSY}$, $SSB < \frac{SSB_{MSY}}{2}$ (i.e., proportion years the stock is overfished), $SSB < 0.3SSB_{F=0}$, $SSB < 0.75SSB_{F=0}$, fully selected $F > F_{MSY}$ (i.e., proportion of years that overfishing occurred), and the proportion of years $Q = 0$ (i.e., proportion of years that the fishery was closed). Interannual variation in yield (IAV) was also recorded:

$$(24) \quad IAV = \frac{\sqrt{\frac{1}{50} \sum_{y=1}^{50} (Y_{y+1} - Y_y)^2}}{\left(\frac{1}{50} \sum_{y=1}^{50} Y_y\right)}$$

Predator performance metrics

Population abundance and recruitment were direct outputs for all modeled predators. Population biomass was directly output for tuna and dogfish. Stakeholders were interested in predator condition for fish and marine mammal predators at the first workshop. While delay difference models do not track individuals or age cohorts, a measure of population average weight (population biomass/population numbers) was output for tuna and dogfish.

Productivity, the number of fledglings per breeding pair, was output for the tern model. Productivity was calculated as adult recruitment times 10 (to account for the 10% survival rate of fledglings to adults) divided by tern abundance 4 years earlier in the simulation.

Stakeholders were interested in different measures of population status depending on the predator. For commercially fished species, status relative to current management reference points was preferred. Tuna and dogfish biomass was divided by a biomass reference point specified in current stock assessments: tuna SSB_{MSY}^p was 13 226 (ICCAT 2015), and dogfish SSB_{MSY}^p was 159 288 (Rago and Sosebee 2010). Because dogfish were fully exploited in our model, they did not reach SSB_{MSY}^p , so we also evaluated status relative to $0.5SSB_{MSY}^p$ (i.e., overfished). Tuna condition status was assessed by dividing the output population average weight with the equilibrium average weight. Common tern colonies are managed to improve productivity, so stakeholders suggested that a common tern productivity level of 0.8 would be a minimum threshold, while a productivity of 1.0 would be a target. In addition, total population status was measured relative to current population numbers using the rationale that maintaining at least the current population was desirable. The average common tern population of nesting pairs (including Monomoy) from 1998 to 2015 was 16 000.

Evaluating the frequency of desirable or undesirable states over the course of a simulation is suggested by Punt et al. (2016a). We calculated two metrics for each of the status determinations. First,

we calculated the minimum number of years in any individual simulation that a metric was above a given threshold. This is a “worst case scenario” metric. Second, we calculated the median proportion of years across all simulations for a control rule that were above the threshold. This is an “average performance” metric addressing how often good status is maintained.

For all metrics other than “frequency of good status” metrics, we report the median value for each simulation. Then, the 25th percentile, the median, and the 75th percentile of these 100 medians were calculated to represent the performance metric for a particular control rule. Results reported here focus on the median.

Economic performance metrics

Median GR and NR were performance metrics that were constructed directly from the economic submodel analogously to the way predator and herring performance metrics were constructed. Stakeholders were also interested in understanding stability of the herring industry. We introduce a new stability performance metric that characterizes the time series of NR as in equilibrium or disequilibrium (Dickey and Fuller 1979). For each simulation, we perform an econometric test of stationarity (Dickey and Fuller 1979) by estimating

$$(25) \quad \Delta NR_t = \beta NR_{t-1} + \xi_1 \Delta NR_{t-1} + \varepsilon$$

Statistical evidence that $\beta = 0$ is evidence of disequilibrium (nonstationarity) of NR while statistical rejections of $H_0: \beta = 0$ in favor of $H_A: \beta < 0$ is evidence of equilibrium (stationarity). Some of the control rules set quotas that are constant for 3 or 5 years; for these policies, we aggregated NR into 3- or 5-year blocks to examine the equilibrium properties across those blocks. The results of these tests are summarized in two ways. First, we use the unweighted Z-transform method from the meta-analysis literature to combine the results of these simulations (Stouffer et al. 1949; Whitlock 2005). This allows for a test of the null hypothesis that a particular control rule implemented on a particular operating model does not produce a stable equilibrium. Defining ϕ as the standard normal cumulative distribution function, we construct

$$(26) \quad Z = \frac{\sum_{i=1}^k \phi^{-1}(1 - p_i)}{\sqrt{(k)}}$$

Z has a standard normal distribution under the null hypothesis of disequilibrium. We define the performance metric Equil1 as the p value associated with rejecting the null hypothesis that particular control rule leads to disequilibrium of NR. Small values of Equil1 are evidence of a stable equilibrium. As a robustness check, we define the performance metric Equil2 as the percentage of simulations in which we reject the null $H_0: \beta = 0$ from eq. 25. Large values of Equil2 are evidence of a stable equilibrium.

Results

Some performance metrics were redundant, showed similar trade-offs, or were expected not to vary among operating models or control rules. For example, the proportion of age-1 herring in the population was insensitive to control rules because the fishery does not select age-1 fish. Such metrics were listed above to document accurately the outcome of the stakeholder process, but results below were focused on metrics likely of broad interest or with possible sensitivities to operating models or control rules. A subset of 11 metrics is presented with abbreviations for reference in figures: herring relative yield (relyield); $\frac{yield}{MSY}$; herring interannual variation in yield (yieldvar); IAV above; the proportion of years $Q = 0$ the herring fishery was closed: (closure); herring rela-

tive spawning stock biomass (relSSB): $\frac{SSB}{SSB_{F=0}}$; the proportion of years the herring stock is overfished (overfished): $SSB < \frac{SSB_{MSY}}{2}$; probability of good tern productivity (ternprod): the median proportion of years that tern fledglings/breeding pair ≥ 1.0 ; probability of good dogfish status (dogstatus): the median proportion of years that dogfish SSB relative $0.5 SSB_{MSY}^p \geq 1.0$; probability of good tuna weight status (tunawt): the median proportion of years that tuna population weight (population biomass/population numbers) \geq average; median herring fishery net revenue (net revenue): as described above; herring fishery stable equilibrium (equil1) if metric is small: as described above; and herring fishery stable equilibrium (equil2) if metric is large: as described above.

Effect of operating models

Several performance metrics performed similarly among operating models and were stable among control rules (Fig. 10). Herring metrics reported relative to reference points were generally less sensitive to operating models than metrics in absolute units (Fig. 10). Consequently, results below focus on herring metrics reported relative to reference points, with the understanding that the operating models differ in meaningful ways (e.g., different MSY) if metrics are reported in absolute units. Dogfish performance metrics were also robust to the operating models and control rules (Fig. 10) and were not considered further. Tuna were affected by variation in herring growth among operating models, but did not vary among control rules (Fig. 10) and were also not considered further because herring growth is not within the control of a management strategy (i.e., HCR). Other than tuna, most metrics were affected by differences in herring M and steepness more than herring growth (Fig. 10). The Equil1 and Equil2 metrics are generally consistent with each other, and indicate that a stable long-run equilibrium of NR is possible for any operating model (Fig. 10).

Effect of HCRs

The constant catch, conditional constant catch, and the biomass-based control rule with a 15% restriction, generally had less inter-annual variation in yield than the other biomass-based control rules (Fig. 11). This stability, however, came at the cost of foregone yield, with fewer options for those three classes of control rule that could achieve yields near MSY (Fig. 11). The constant catch and biomass-based control rule with a 15% restriction also had more alternatives that led to poorer tern production than other control rules (Fig. 11). The conditional constant catch rule, however, performed well for tern production because the cap of $0.5F_{MSY}$ on fishing mortality effectively prevented levels of mortality that decreased herring abundance to levels where tern production would be compromised. Figure 11 shows that the biomass-based control rule with a 15% restriction would often lead to disequilibria of NR and the HCR type responsible for the long right tails seen in Fig. 10. This finding of a disequilibrium is likely occurring because the 15% restriction explicitly introduces memory into the HCR, causing current period NR to be related to previous period landings. The disequilibrium tests applied to HCRs that have 3- or 5-year constant quotas use less data (the 10- or 16-year time periods corresponding to the final 50 years) and therefore have lower statistical power than those applied to the HCRs that are set annually. However, the standard BB3yr and BB5yr HCRs are characterized by equilibrium (Fig. 11), so we are doubtful that this finding is caused by a statistical power problem. Ultimately, the NEFMC eliminated the constant catch, conditional constant catch, and biomass-based rule with a 15% restriction from consideration because of the foregone yield and (or) relatively poor performance of some alternatives for the tern production metric.

Given that the metrics reported in results below focus on those that were relatively robust to operating models, additional results

were only reported for the operating model with low M , high steepness, and slow growth ("HighSlowCorrect" in figures). Further, due to NEFMC decisions to eliminate the constant catch, conditional constant catch, and a biomass-based rule with a 15% restriction, more detailed trade-off analyses below focus on the biomass-based control rule applied annually (BB), or with 3- (BB3yr) or 5-year (BB5yr) quota blocks.

Trade-offs

At similar levels of yield, using 3- or 5-year blocks for the biomass-based control rule produced more control rule shapes with less SSB, such that the short-term stability of such quota blocks comes at the potential cost of less SSB (Fig. 12). All three biomass-based variants had alternatives that could achieve yield near MSY over a range of SSB levels.

At similar levels of yield, the application of quota blocks resulted in lower extreme highs in IAV (Fig. 13). The application of quota blocks also reduced the number of alternatives near MSY, further illustrating the trade-off between stability and yield.

All three biomass-based alternatives could achieve $>90\%$ MSY with nearly zero fishery closures, although the number of alternatives was fewer with longer quota blocks (Fig. 14). While the trade-offs were generally similar between $\frac{yield}{MSY}$ and the frequency that $Q = 0$ for all three alternatives, vertical patterning becomes evident with longer quota blocks. This patterning was caused by biomass-based control rule shapes becoming more alike with longer quota blocks, with their performance dominated more by the length of fishery closures dictated by the quota block length than by the specific shape of the control rule. As an extreme example, if a 50-year quota block was applied, any control rule that would close the fishery early in the time series would have a nearly identical frequency of $Q = 0$ near 100%, but these same control rules would behave much differently with shorter quota blocks where the responsiveness of the specific control rules would drive results.

All of the control rule alternatives offered options that had near zero frequency of $SSB < \frac{SSB_{MSY}}{2}$ (Fig. 15). At similar levels of yield, using 3- or 5-year blocks for the biomass-based control rule produced more control rule shapes with higher frequency of $SSB < \frac{SSB_{MSY}}{2}$, such that the short-term stability of such quota blocks comes at the potential cost of more frequently dropping to relatively low levels of biomass.

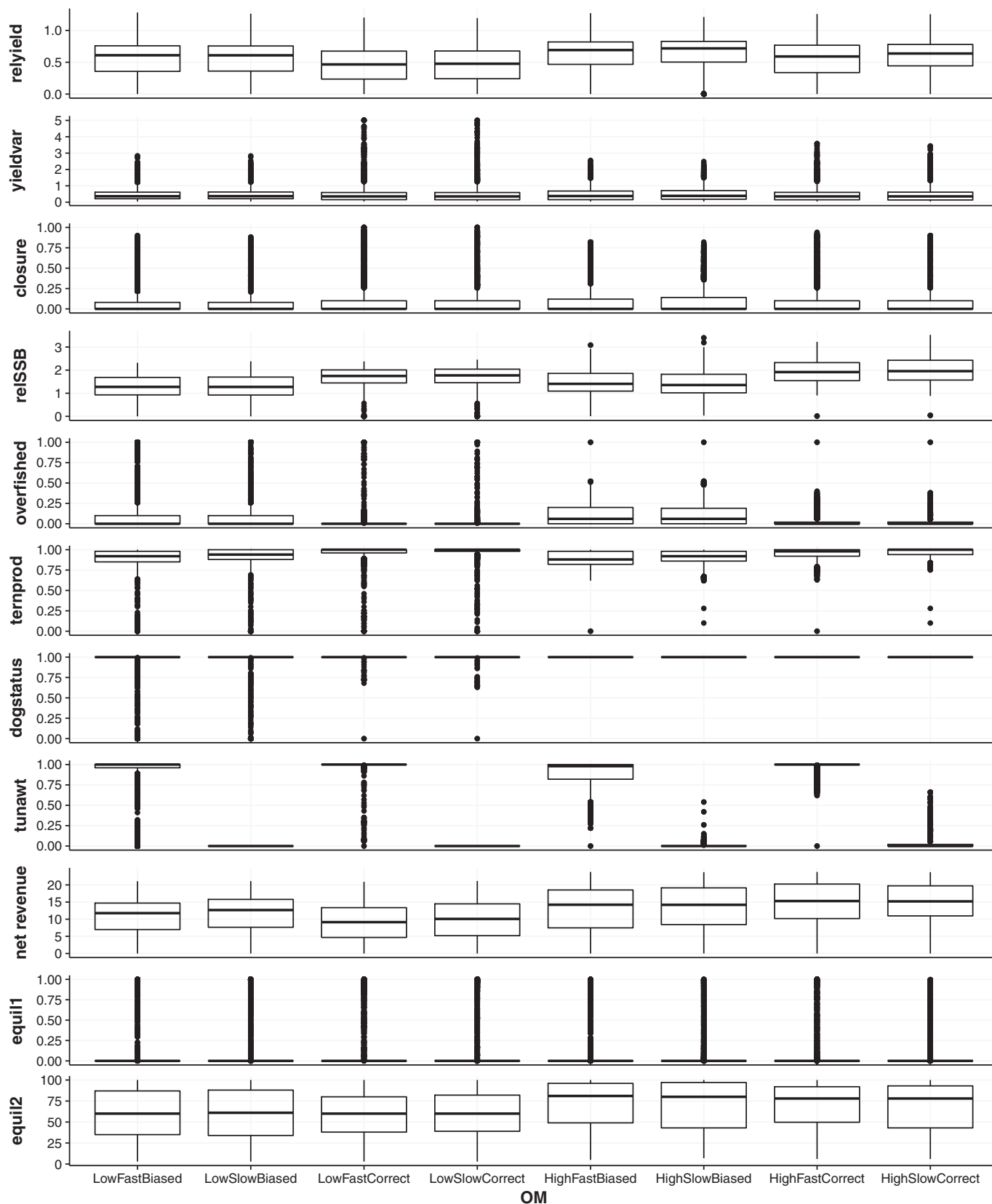
All of the control rule alternatives offered options that had high frequency of good tern productivity (equal to or greater than the management target of 1.0; Fig. 16). At similar levels of yield, using 3- or 5-year blocks for the biomass-based control rule produced more control rule shapes with slightly lower frequency of good tern productivity; however, the difference is between $>80\%$ and $>90\%$ frequency of good productivity.

The control rule alternatives resulted in good stability characteristics across a range of NR for the fishery (Fig. 17). We have included results from the biomass-based control rule with a 15% restriction on interannual variation in the quota to illustrate the relatively poor performance of this control rule type.

Discussion

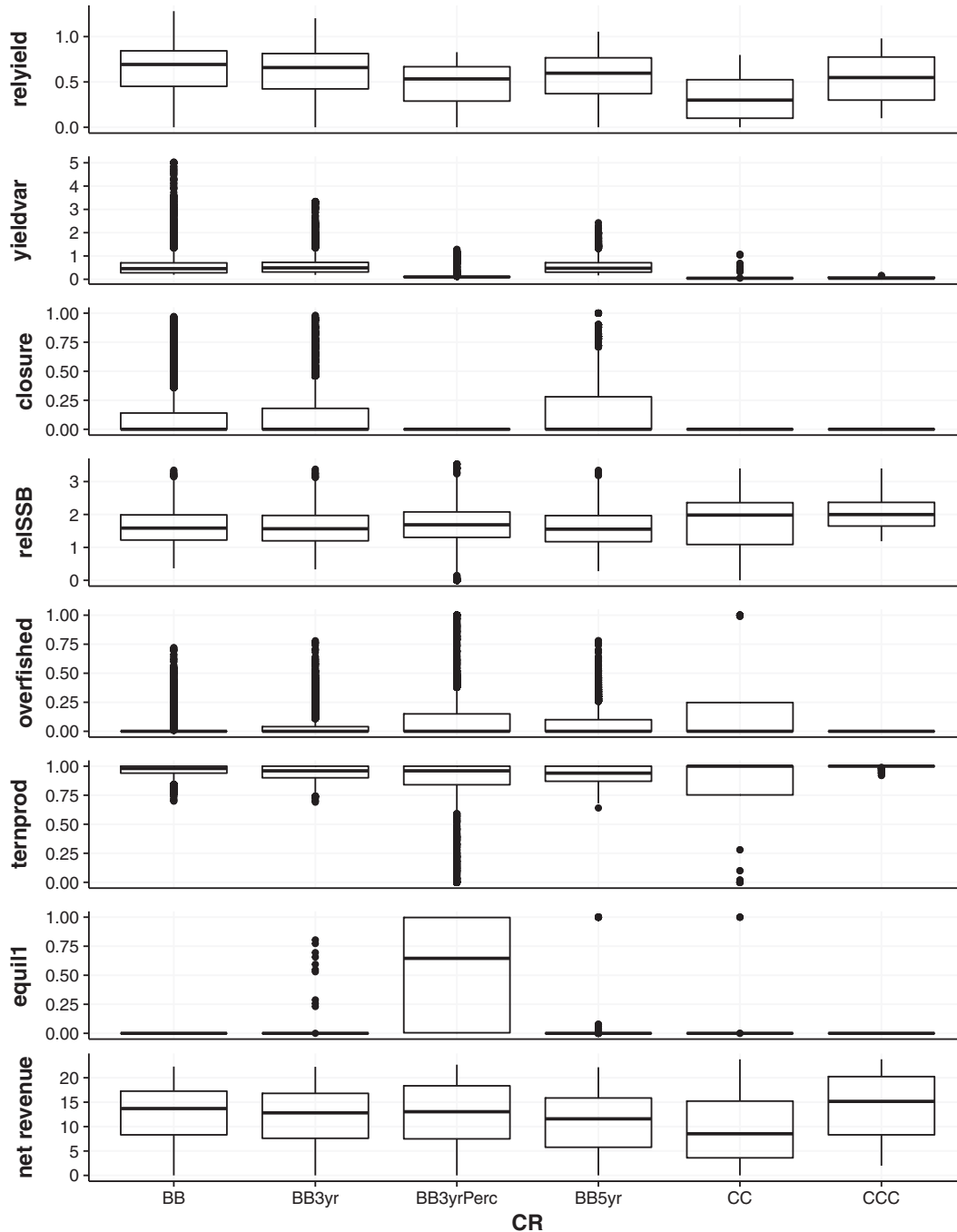
Results were generally robust among operating models when presented relative to biological reference points (e.g., $\frac{yield}{MSY}$), but absolute scale differed among operating models. This result has been previously reported (Deroba and Bence 2012) and is likely why the results of most MSEs are presented in relative units (A'mar et al. 2009; Deroba and Bence 2012; Wiedenmann et al. 2017). While convenient scientifically, the differences in absolute scale can create challenges when communicating with stakeholders because most people are more comfortable with absolute

Fig. 10. Differences between operating models (OM) for key metrics. Boxplots represent the median (wide line within the box) and 25th and 75th percentiles (box) of the distribution of medians for final 50 years of each simulation for each performance metric (x-axis label; see text for definitions) and operating model (see Table 2 for definitions) across all control rule types. Boxplot whiskers include the highest and lowest observations within 1.5 box lengths from the box. Observations further than 1.5 box lengths from the box (outliers) are represented by points.



Can. J. Fish. Aquat. Sci. Downloaded from cdsciencepub.com by NOAA CENTRAL on 02/29/24
For personal use only.

Fig. 11. Differences between control rule types for key metrics. Boxplots represent the median (wide line within the box) and 25th and 75th percentiles (box) of the distribution of medians for final 50 years of each simulation for each performance metric (x-axis label; see text for definitions) and control rule type (CR; described in the text Harvest control rules section: biomass-based with 1-year (BB), 3-year (BB3yr), and 5-year (BB5yr) quota blocks, biomass-based 3-year quota block with a 15% restriction on interannual quota change (BB3yrPerc), constant catch (CC), and conditional constant catch(CCC)) across all operating models. Boxplot whiskers include the highest and lowest observations within 1.5 box lengths from the box. Observations further than 1.5 box lengths from the box (outliers) are represented by points.



units (e.g., fishing industry representatives think in tons and not fractions of MSY; Feeney et al. 2019). Thus, reporting results in absolute units will still have value in public settings, and scientists should attempt to convey the reasons for differences in scale among operating models (Feeney et al. 2019).

The constant catch, conditional constant catch, and restricting annual changes in the quota by 15% produced less variable yield than the other biomass-based alternatives, but at the expense of yield, more frequent low levels of herring biomass, and more

outcomes relatively detrimental to predators. This result is consistent with previous simulations of similar HCRs for lake whitefish (*Coregonus clupeaformis*) in the Laurentian Great Lakes (Deroba and Bence 2012) and a roundfish stock managed by the International Council for the Exploration of the Sea (Kell et al. 2006). For lake whitefish; Deroba and Bence (2012) found that a conditional constant catch rule could achieve similar stability in yield as a 15% restriction on the interannual change to a quota applied to a biomass-based control rule, but with higher yields. A broader

Fig. 12. Trade-off between herring relative yield and relative SSB for HighSlowCorrect operating model and biomass-based control rules with 1-year (BB), 3-year (BB3yr), or 5-year (BB5yr) quota blocks. Each point represents the median of 100 medians taken over the final 50 years of each simulation. [Colour online.]

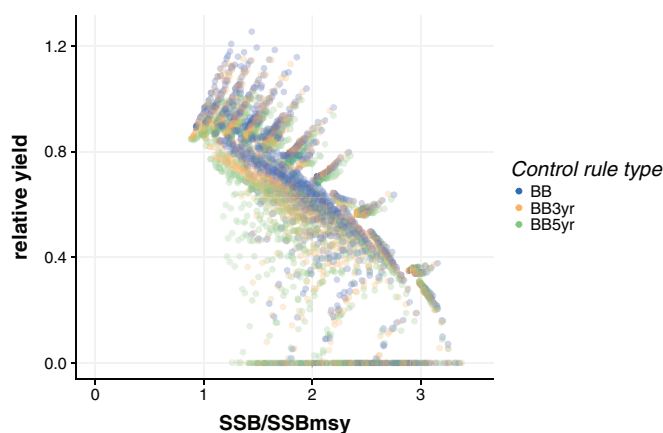
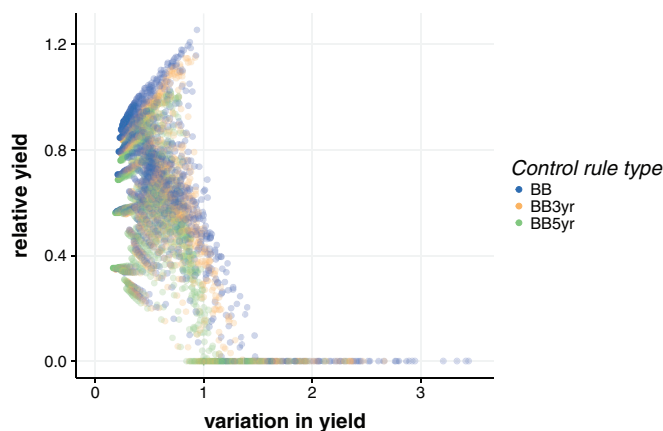


Fig. 13. Trade-off between herring relative yield and variation in yield for HighSlowCorrect operating model and biomass-based control rules with 1-year (BB), 3-year (BB3yr), or 5-year (BB5yr) quota blocks. Each point represents the median of 100 medians taken over the final 50 years of each simulation. [Colour online.]



range of percent restrictions on the interannual variation in the quota could be evaluated, however, before assuming that these conclusions are general. A 15% restriction was used in this analysis because that amount was specified as desirable by stakeholders. Conceivably, a different percentage may strike a more agreeable trade-off between stability and other metrics of interest. The performance of restraints on the interannual variation in quotas also depends on stock status and variation in life-history traits, such as growth (Punt et al. 2002; Kell et al. 2006), and so should likely be evaluated on a case-by-case basis.

If short-term stability in yield is a fishery objective, then using quota blocks where the target harvest is the same for multiple years may be an effective method that costs little in the performance of other metrics. The biomass based options with 3- or 5-year quota blocks produced similar ranges of performance and similar trade-offs as annual changes in the quota. Thus, short-term stability could be gained at little long-term cost with appropriately selected harvest policy parameters. Although, at similar level of yields, the quota block alternatives produced more options with lower biomass. Thus, a closed-loop simulation could be used to evaluate the changes in relative performance between applying a control rule annually or using multiyear specifications.

Fig. 14. Trade-off between herring relative yield and frequency of fishery closure for HighSlowCorrect operating model and biomass-based control rules with 1-year (BB), 3-year (BB3yr), or 5-year (BB5yr) quota blocks. Each point represents the median of 100 medians taken over the final 50 years of each simulation. [Colour online.]

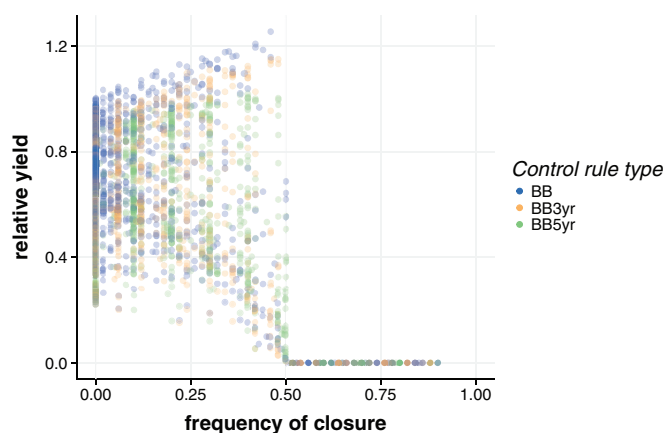
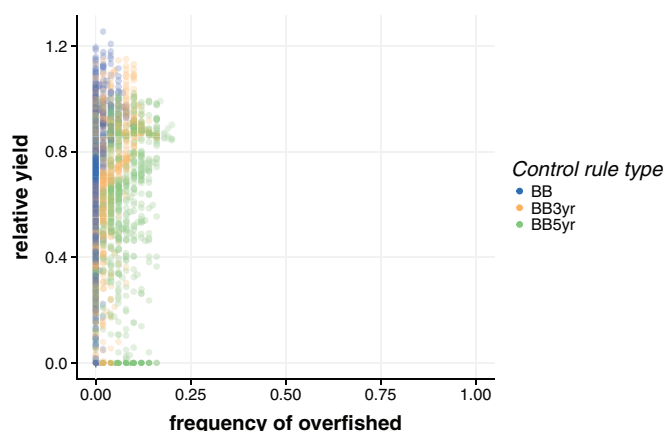


Fig. 15. Trade-off between herring relative yield and probability of the stock being overfished for HighSlowCorrect operating model and biomass-based control rules with 1-year (BB), 3-year (BB3yr), or 5-year (BB5yr) quota blocks. Each point represents the median of 100 medians taken over the final 50 years of each simulation. [Colour online.]

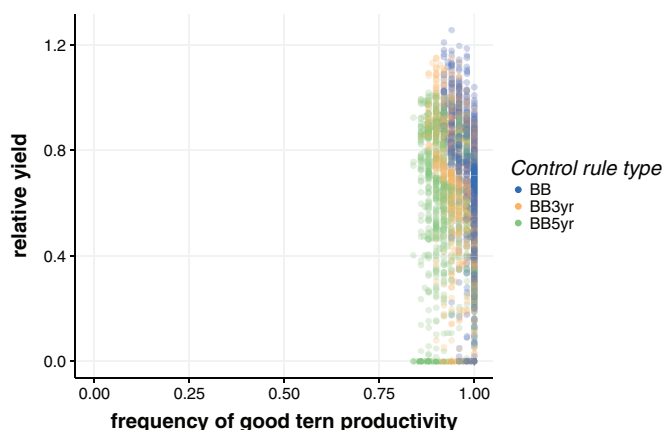


In a simulation of lake whitefish, Li et al. (2016) found that the method of specifying target harvests between stock assessments was less important to relative performance than assessment frequency. Furthermore, setting target harvest to the same value for multiple years between assessments, as in the quota blocks used here, performed similarly to using projections in the interim years.

Nearly every stakeholder group proposed metrics that measured the health of the herring stock. The herring biomass performance metric could capture passive (non-use) value associated with herring biomass. It may also be related to active use values if higher herring biomass improves ecotourism, recreational, or commercial fishery outcomes. Finally, a herring biomass performance metric may be a proxy for unmodeled recreational, ecotourism, or socioeconomic components of the ecosystem that are believed to be positively associated with herring biomass.

The simulation model used in this MSE can be thought of as a model of intermediate complexity (Plagányi et al. 2014). Future MSE models that are designed to address the role of herring as forage in the ecosystem should graduate to advanced complexity. These advances should include more realistic models of the scien-

Fig. 16. Trade-off between herring relative yield and probability of good tern productivity for HighSlowCorrect operating model and biomass-based control rules with 1-year (BB), 3-year (BB3yr), or 5-year (BB5yr) quota blocks. Each point represents the median of 100 medians taken over the final 50 years of each simulation. [Colour online.]



tific process, herring stocks, predator–prey relationships, human behavior, and ecosystem valuation methods. However, linking the appropriate components to derive the desired set of performance metrics for management decision-making need not be an insurmountable challenge. For example, fairly simple methods were used to link environmental drivers, prey and predator species, and detailed social and economic components of fisheries in the Gulf of Alaska (Zador et al. 2017). The expert opinion of stakeholders involved in the MSE can be used to specify similar conceptual models in New England. In future iterations of the MSE, conceptual and qualitative modeling might be used to map out where critical data and further integrated model development would most efficiently address priority management needs.

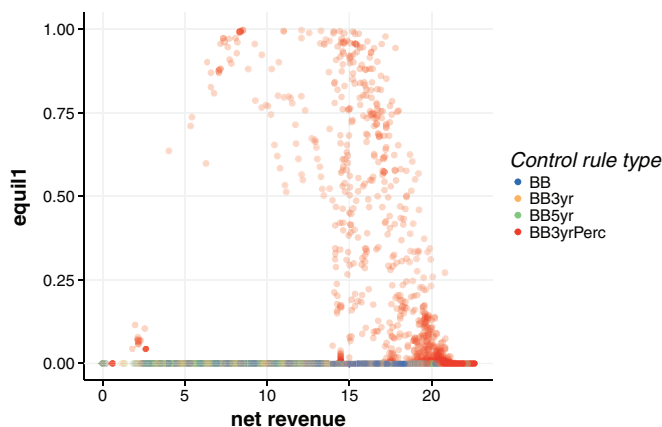
The dream

In this section, we discuss the limitations of the models and simplifications necessary to meet decision-making time frames while incorporating ecosystem and economic models into management strategy evaluation.

Herring dreams

The application of the biomass-based and conditional constant catch control rules assumed MSY reference points were known without error, as in several other studies (Irwin et al. 2008; Punt et al. 2008). In reality, such reference points are likely to be uncertain in most cases, and the bias and precision of reference point estimates depend on life-history traits and autocorrelation in recruitment (Haltuch et al. 2008, 2009). How the relative performance of control rules would change in the presence of error in reference points is unclear and should be a topic for future research. Given that assessment error affects relative control rule performance (Deroba and Bence 2008), errors in reference point estimation may compound those issues and affect results in meaningful ways. Incorporating realistic errors in reference point estimation, however, will be challenging because the effects will depend on exploitation history and the degree to which life-history traits vary among years (Brodziak et al. 2008; Legault and Palmer 2016). This MSE evaluated uncertainty in life-history traits, but not time variation in those traits. Methods for estimating reference points and the proper response of management in the presence of time-varying life-history traits is an active area of research and has been a focal point of MSEs (A'mar et al. 2009; Legault and Palmer 2016). For example, increases in natural mortality through time have different implications on reference

Fig. 17. Trade-off between herring fleet net revenue and Equil1 for HighSlowCorrect operating model and biomass-based control rules with 1-year (BB), 3-year (BB3yr), or 5-year (BB5yr) quota blocks compared with 3-year quota block with a 15% restriction on annual quota changes (BB3yrPerc). Each point represents the median of 100 medians taken over the final 50 years of each simulation. [Colour online.]



points depending on whether a per recruit approach is used or if management is concerned with limiting total mortality, where a per recruit approach would increase target F but limiting total mortality would require decreasing target F (Legault and Palmer 2016). An MSE for Gulf of Alaska walleye pollock (*Gadus chalcogrammus*) demonstrated that management strategies (i.e., the combination of HCR and estimation model) performed differently depending on how reference points were estimated in the presence of regime shifts in recruitment, with estimates of unfished biomass depending on the range of years used to estimate average recruitment (A'mar et al. 2009). Atlantic herring have experienced time-varying growth, and the possibility of time-varying natural mortality has been considered an uncertainty in recent stock assessments (Northeast Fisheries Science Center 2012; Deroba 2015). Thus, evaluating the effect of errors in reference point estimation would be a prudent advancement of the MSE. The stakeholder-driven process, however, was already overwhelming for many participants given the time frame, and so adding this additional realism would likely be best completed as a separate exercise that would rely more on technical experts than stakeholder input.

This herring MSE did not incorporate a true stock assessment as part of the management strategy. The approach taken offered the advantage of simplicity, brevity, and the ability to control bias in the assessment without having to specify a source of the bias (e.g., misspecified M , incorrect selectivity), as would be required with the incorporation of a true stock assessment. This MSE also took advantage of pre-existing and quality-checked code that helped meet the decision-making time frames. Amending this code to include a true stock assessment likely would have precluded attaining the deadlines. Conceivably, however, relative control rule performance could vary depending on the misspecification that induces a bias in a true stock assessment model, and so several alternative sources of bias would likely have to be evaluated as part of an MSE, substantially increasing computing time in an already demanding time frame. Given the technical nature of stock assessments, constructing an evaluation of the effect of different sources of bias in a stock assessment also could likely be conducted outside of a stakeholder process. Nonetheless, incorporating a true stock assessment is considered best practice and may induce realism, such as levels of autocorrelation among assessment fits, that cannot be reproduced otherwise (Cox and Kronlund 2008; Punt et al. 2016a). Consequently, this MSE could

suggest alternative management decisions as such additional realism is implemented.

Predator dreams

As has been found in other MSE analyses (Punt et al. 2016a), the predator results may be more useful for eliminating poor control rule options (BB3yrPerc, CC) than for optimizing herring control rules to improve predator metrics. There are several reasons for this. Predator populations are affected by many factors, while we attempted to isolate factors associated with prey dynamics. Further, in the Northeast US, predators have many prey options (Link 2002), while we attempted to evaluate relationships with just one prey, herring. Finally, time limitation necessitated simple, tractable models of complex ecological relationships. Our approach was to use the best-supported relationship for each predator based on observations from the Northeast US ecosystem. We discuss the pros and cons of this approach for each predator below.

Western Atlantic bluefin tuna migrate widely and forage throughout the North Atlantic; their population footprint is much larger than that of Northeast US Atlantic herring. However, tuna feed seasonally in the Gulf of Maine, exploiting concentrated, high-energy prey to maximize growth (Golet et al. 2013). Because tuna growth is key in the Northeast US, and because there is a well-supported relationship between herring weight and tuna growth (Golet et al. 2015), we used this relationship. Other relationships were also investigated. Available data do not suggest a positive relationship between herring and tuna populations in our models for this MSE; Northeast US shelf herring have increased during a period of bluefin tuna decline (Northeast Fisheries Science Center 2012; ICCAT 2015). Stakeholder observations and fine-scale analyses (e.g., Golet et al. 2013) suggest that bluefin tuna follow herring in the Gulf of Maine and likely aggregate around herring while feeding. However, our models designed to address acceptable biological catch control rules at the Northeast US shelf scale do not address herring–tuna interactions in a specific place or time, and we can draw no conclusions from our modeling about predator–prey co-occurrence or availability at smaller, local scales. Similarly, without additional observations, extrapolating local-scale co-occurrence to population-level relationships is not well supported. Future iterations of the MSE would require finer-scale data on predators, prey, and fisheries for both to address these questions.

Common terns, in contrast, are seasonal central-placed foragers near their island breeding colonies in the Gulf of Maine (Nisbet 2002). Their foraging footprint during chick production season is much smaller than the scale of the Northeast US Atlantic herring population. Because tern productivity is a key management objective for tern colonies in the Gulf of Maine, we used the substantial existing data to explore a relationship between herring populations and tern reproductive success. However, many factors other than herring abundance affect tern production. According to Gulf of Maine Seabird Working Group minutes and other work, predation by mammals, gulls, and other birds is a major factor that most colony management aims to control (Donehower et al. 2007; Scopel and Diamond 2017b). Further, timing of weather events and prey availability is important for production, but difficult to quantify at all colonies from current data (Scopel and Diamond 2017a). Similarly, the relatively small spatial scale and depth distribution of prey affects tern foraging success as well as the overall abundance of prey (Scopel et al. 2018). At one colony during the same year, the proportion of herring in tern chick diets was much lower than the proportion of herring in razorbill (*Alca torda*) diets at the same colony; razorbills are capable of deeper dives than terns (GOMSWG minutes). Spatial variability of predation, weather, and prey distribution may drive the high variation in observed herring population – tern productivity relationship among colonies, similar to observations in other ecosystems (Sydeman et al. 2015). This high variance in the observations is not considered by

the modeled herring–tern relationship. Further, the tern model is optimistic about population trajectory because it considers only herring total biomass effects on terns and does not model predation, habitat quantity and quality, etc. In future iterations of the MSE, a better match of spatial scale for fishery removals and seabird foraging (similar to tunas above) could better investigate management options that are not easily addressed with a stock-scale annual HCR (Sydeman et al. 2017).

Spiny dogfish may have the best spatial footprint match with Atlantic herring in the Northeast US of the three predators modeled. Dogfish forage through the same range as herring for most of the year. Considerable information on dogfish diet has been collected over time in the region, and there are adequate data to conduct a stock assessment. However, the dogfish relationship assumes herring abundance improves dogfish survival because no clear relationship was found with recruitment or growth. Increased survival may not be the mechanism for the observed positive influence of herring in diet on the dogfish population.

Our approach allowed “bottom-up” effects of herring on predators to be examined, which was the key management question. Although we selected predators with high herring diet proportions, observed predator population responses to herring alone do not dominate dynamics, and our herring–predator relationship models reflect that. Predator responses to aggregate prey dynamics may be much clearer than responses to individual prey in the Northeast US ecosystem given its food web structure with many alternative prey (Link 2002). Further, food web modeling explored here and in other studies suggested that “top-down” effects of predators on herring, simultaneous interactions of multiple predators with herring, and side effects on other forage species could be important in this ecosystem (Link et al. 2006, 2008, 2009). While modeling multispecies interactions is a more complex and time-consuming undertaking, the results may give clearer advice for managers making decisions regarding multiple simultaneously exploited prey and predators within the ecosystem (DeWitt and Langerhans 2003; Lovvorn et al. 2013).

Here, the general objective for the NEFMC was to answer “how do changes in herring population abundance affect predator populations?” This is a different and more complex question than that addressed in the 2012 herring assessment: “how much herring is consumed by predators?” Our MICE models were designed specifically for evaluating alternative herring control rules, not predator stock assessment and population prediction. NEFMC specifications and time constraints did not permit development of integrated multispecies models addressing both bottom-up herring impacts on predators and predation mortality on herring, nor permit spatial or seasonal models accounting for migrations of wide-ranging predators into or out of the Northeast US shelf ecosystem. Existing multispecies models in the region only account for predation mortality on herring (Curti et al. 2013; Gaichas et al. 2017). A MICE approach could be taken using these models as a starting point to incorporate key multispecies feedbacks and even broadly spatial interactions in future iterations of the MSE, as was done for the California Current (Punt et al. 2016b).

We caution against generalizing results for these particular predators to other predators, as population parameters and herring relationships differ. Although considerable work has been done examining forage fish fishing in many ecosystems (Cury et al. 2011; Essington et al. 2015; Hilborn et al. 2017), and it can be tempting to generalize control rules specific to forage fish across all ecosystems (Pikitch et al. 2012), our results demonstrated that many potential control rules for Northeast US herring gave equally good results for the modeled herring predators. While we do not suggest that relationships we found here will hold for predators and prey in other ecosystems, we wholeheartedly recommend the use of ecosystem-specific data to evaluate forage fish HCRs and trade-offs between objectives on a case-by-case basis.

Economic dreams

The economic model of the herring fishery did not include fixed costs. If firms do not enter or exit, then the exclusion of fixed costs from the model has minimal effect when comparing HCRs within or across OMs. NR would all be overestimated by the same fixed amount. The stationarity metric would be unaffected; however, IAV constructed without fixed costs will be smaller than the true IAV that contains fixed costs. In reality, firms can enter and exit this industry. Economic theory suggests that firms will enter (exit) if they anticipate large positive (negative) profits over a particular planning horizon. While herring is a limited-access fishery, less than three-quarters of the permits vessels are active, suggesting that firms could enter. Understanding exactly how these entry and exit decisions are made was not possible on the timeline requested by NEFMC.

We also assume that marginal costs are equal to average variable costs, constant for each fleet, and do not depend on the level of biomass. A more rigorous approach might include estimating a (economic) production function for the herring fishery; this was not done in the interest of time. It is difficult to predict how estimating a true cost function and integrating those results would change the results of the study.

Catch in the economic model can be different from both quota and yield (from the herring model). This is frequently handled as symmetric implementation error in the fisheries literature (as it is in this model). However, the economic model suggests that either an asymmetric error term, in which the error depends on prices of inputs and outputs, or a more integrated biological and economic model is warranted.

Consumer welfare measures could be determined from a demand curve for herring. Equation 22 estimates a price–quantity relationship, which is not necessarily a demand curve for herring. Rigorous estimation of a demand curve for herring requires modeling all goods that are substitutes for herring, including mackerel, menhaden, squid, and other substitute baits. This was not done due to limited time available. Kirkley et al. (2011) use a static input–output model to simulate the effects of changes in herring quotas and predator biomass levels on the New England economy. Because the Kirkley et al. (2011) analysis suggested that the effects of changes in herring catch on other segments of the economy are quite small, economic analysis was confined to the herring fleet.

Economic methods that can inform ecosystem approaches to fisheries management include portfolio methods (Edwards et al. 2004; Jin et al. 2016), coupled ecosystem–region models (Jin et al. 2003, 2012; Kirkley et al. 2011), and bioeconomic models (Tschirhart 2000; Finnoff and Tschirhart 2003; Brown et al. 2005; Lehuta et al. 2014) of varying complexity. However, there was simply not enough time to employ these methods, nor to link economic outcomes to sociocultural outcomes (Zador et al. 2017). The largest limitation of the economic model is that only the herring fishery is quantitatively modeled. Humans who indirectly use in situ herring were not formally modeled.

The predator section models a few representative consumers of live herring: terns, tuna, whales, and predatory fish. Ecosystem valuation methods could be used to measure changes in outcomes for those species in dollar value (Loomis and White 1996; Richardson and Loomis 2009; Lew 2015). People may derive value from changes in the status of these predators through either use or non-use values. For example, people may directly value higher abundances of an animal or protection of an endangered species, even if they have no plans to watch or view them (Lew et al. 2010; Lew and Wallmo 2017). Quantifying these values typically is done using stated preference methods with data collected using surveys; these studies are costly and time-consuming to develop and conduct rigorously. Benefit transfer, a method in which valuation from previous studies is applied to a new study area, may be a way to overcome these barriers (Navrud and Ready 2007; Johnston and Rosenberger 2010).

Stakeholders may also derive use value from changes in the modeled species. Changes in herring biomass may change costs, catches, or prices in the commercial fishery for a predator. For example, increases in biomass of spiny dogfish could lead to both higher quotas and lower costs to catch more abundant fish. Examining changes in costs would require an economic model of production for the spiny dogfish fishery (similar to the model not used for the herring fleet; Holland and Sutinen 1999; Hutniczak 2014; Reimer et al. 2017). Changes in product quality could affect prices (Larkin and Sylvia 1999; Asche et al. 2015). For example, because larger tuna receive higher prices, the effects of changes in average weight could be deduced from existing hedonic models (McConnell and Strand 2000; Carroll et al. 2001).

These changes are not confined to an extractive sector of the economy. For example, changes in whale populations may change outcomes for whale-watching customers. This type of value could be quantified using both stated and revealed preference data, both of which typically require collection of survey data (Larson et al. 2004). These studies are often quite costly and time-consuming to develop and conduct rigorously. Perhaps more discouragingly, the precise good being valued needs to be known quite early in the research process. In this application, development of a valuation survey for terns would not have been able to begin until after the predator modeling was nearly complete. Benefit transfer may be the only way to value some of the use values on the time frame required by NEFMC.

Note that increases in the biomass of a particular predator could be a net “bad” for society. For example, an increase in the biomass of a predator that is low-valued but skilled at consuming herring could result in disproportionate increases in that low-valued predator. If that low-valued predator is not a complete specialist (in consuming herring), it may also drive down the biomass of high-valued predators. The ability to manipulate the ecosystem with a prey-level acceptable biological catch control rule to achieve desirable outcomes depends on the rates at which these increases in prey are converted into social utility. This conversion depends on the ecosystem technology (conversion of prey into additional biomass of high- and low-valued predators), human technology (conversion of prey and predator biomass into catch or tourism), and human preferences (converting catch or tourism into utility). Despite our current efforts, many of these relationships are not particularly well understood at this time.

To meet management timelines, the herring, predator, and economic models were developed in parallel and not in sequence. This required an educated guess about which predators and predator outcomes were sensitive to the range of HCRs. For example, prior to completing the predator models, we did not anticipate that tuna weight would be sensitive to various operating models (i.e., fast and slow herring growth) but not the HCRs. Developing the herring, predator, and economic models sequentially could have allowed for a model of the tuna fishery that accounts for the size-dependent prices (Carroll et al. 2001). In contrast, devoting scarce research time to examine the costs of harvesting tuna would not have resulted in performance measures that would help managers select among control rules because changes in herring abundance had relatively little effect on tuna.

Acknowledgements

We thank the more than 80 individuals from the fishing industry, federal and state agencies, academia, environmental organizations, and other entities who gave their time to contribute to the workshops. Madeleine Hall-Arber, Jessica Joyce, Laura Singer, and Tiffany Vidal provided excellent workshop facilitation. Beth Goettel, Linda Welch, and Aly McKnight contributed seabird data and expertise. This manuscript benefitted from the peer review panel: Lisa Kerr (Chair), Gavin Fay, Douglas Lipton, and John Wiedenmann. Statements herein should neither be considered formal positions of the NEFMC or NMFS nor a consensus of the

participants. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government.

References

- A'mar, Z.T., Punt, A.E., and Dorn, M.W. 2009. The impact of regime shifts on the performance of management strategies for the Gulf of Alaska walleye pollock (*Theragra chalcogramma*) fishery. *Can. J. Fish. Aquat. Sci.* **66**(12): 2222–2242. doi:10.1139/F09-142.
- Asche, F., Chen, Y., and Smith, M.D. 2015. Economic incentives to target species and fish size: prices and fine-scale product attributes in Norwegian fisheries. *ICES J. Mar. Sci.* **72**(3): 733–740. doi:10.1093/icesjms/fsu208.
- Brodziak, J., Cadrin, S.X., Legault, C.M., and Murawski, S.A. 2008. Goals and strategies for rebuilding New England groundfish stocks. *Fish. Res.* **94**(3): 355–366. doi:10.1016/j.fishres.2008.03.008.
- Brown, G., Berger, B., and Ikiara, M. 2005. A predator–prey model with an application to Lake Victoria fisheries. *Mar. Resour. Econ.* **20**(3): 221–247. doi:10.1086/mre.20.3.42629473.
- Bubley, W.J., Kneebone, J., Sulikowski, J.A., and Tsang, P.C.W. 2012. Reassessment of spiny dogfish *Squalus acanthias* age and growth using vertebrae and dorsal-fin spines. *J. Fish Biol.* **80**(5): 1300–1319. doi:10.1111/j.1095-8649.2011.03171.x. PMID:22497385.
- Butterworth, D.S. 2007. Why a management procedure approach? Some positives and negatives. *ICES J. Mar. Sci.* **64**(4): 613–617. doi:10.1093/icesjms/fsm003.
- Carroll, M.T., Anderson, J.L., and Martínez-Garmendia, J. 2001. Pricing U.S. North Atlantic bluefin tuna and implications for management. *Agribusiness*, **17**(2): 243–254. doi:10.1002/agr.1014.
- Chase, B.C. 2002. Differences in diet of Atlantic bluefin tuna (*Thunnus thynnus*) at five seasonal feeding grounds on the New England continental shelf. *Fish. Bull.* **100**(2): 168–180. Available from <http://aquaticcommons.org/15201/> [accessed 26 April 2016].
- Clark, W.G., and Hare, S.R. 2004. A conditional constant catch policy for managing the Pacific halibut fishery. *N. Am. J. Fish. Manage.* **24**(1): 106–113. doi:10.1577/M03-020.
- Collie, J.S., Botsford, L.W., Hastings, A., Kaplan, I.C., Largier, J.L., Livingston, P.A., Plagányi, É., Rose, K.A., Wells, B.K., and Werner, F.E. 2016. Ecosystem models for fisheries management: finding the sweet spot. *Fish. Fish.* **17**(1): 101–125. doi:10.1111/faf.12093.
- Cox, S.P., and Kronlund, A.R. 2008. Practical stakeholder-driven harvest policies for groundfish fisheries in British Columbia, Canada. *Fish. Res.* **94**(3): 224–237. doi:10.1016/j.fishres.2008.05.006.
- Curti, K.L., Collie, J.S., Legault, C.M., Link, J.S., and Hilborn, R. 2013. Evaluating the performance of a multispecies statistical catch-at-age model. *Can. J. Fish. Aquat. Sci.* **70**(3): 470–484. doi:10.1139/cjfas-2012-0229.
- Cury, P.M., Boyd, L.L., Bonhommeau, S., Anker-Nilssen, T., Crawford, R.J.M., Furness, R.W., Mills, J.A., Murphy, E.J., Osterblom, H., Paleczny, M., Piatt, J.F., Roux, J.-P., Shannon, L., and Sydeman, W.J. 2011. Global seabird response to forage fish depletion—one-third for the birds. *Science*, **334**(6063): 1703–1706. doi:10.1126/science.1212928. PMID:22194577.
- Deroba, J.J. 2014. Evaluating the consequences of adjusting fish stock assessment estimates of biomass for retrospective patterns using Mohr's Rho. *N. Am. J. Fish. Manage.* **34**(2): 380–390. doi:10.1080/02755947.2014.882452.
- Deroba, J.J. 2015. Atlantic herring operational assessment report 2015. US Dept. of Commerce, Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543, USA.
- Deroba, J.J., and Bence, J.R. 2008. A review of harvest policies: understanding relative performance of control rules. *Fish. Res.* **94**: 210–223. doi:10.1016/j.fishres.2008.01.003.
- Deroba, J.J., and Bence, J.R. 2012. Evaluating harvest control rules for lake whitefish in the Great Lakes: accounting for variable life-history traits. *Fish. Res.* **121–122**: 88–103. doi:10.1016/j.fishres.2012.01.015.
- DeWitt, T.J., and Langerhans, R.B. 2003. Multiple prey traits, multiple predators: Keys to understanding complex community dynamics. *J. Sea Res.* **49**(2): 143–155. doi:10.1016/S1385-1101(02)00220-4.
- Dickey, D.A., and Fuller, W.A. 1979. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **74**(366a): 427–431. doi:10.1080/01621459.1979.10482531.
- Donehower, C.E., Bird, D.M., Hall, C.S., and Kress, S.W. 2007. Effects of gull predation and predator control on tern nesting success at Eastern Egg Rock, Maine. *Waterbirds*, **30**(1): 29–39. doi:10.1675/1524-4695(2007)030[0029:EOGPAP]2.0.CO;2.
- Edwards, S.F., Link, J.S., and Rountree, B.P. 2004. Portfolio management of wild fish stocks. *Ecol. Econ.* **49**: 317–329. doi:10.1016/j.ecolecon.2004.04.002.
- Essington, T.E., Moriarty, P.E., Froehlich, H.E., Hodgson, E.E., Koehn, L.E., Oken, K.L., Siple, M.C., and Stawitz, C.C. 2015. Fishing amplifies forage fish population collapses. *Proc. Natl. Acad. Sci.* **112**(21): 6648–6652. doi:10.1073/pnas.1422020112. PMID:25848018.
- Feeney, R.G., Boelke, D., Deroba, J.J., Gaichas, S.K., Irwin, B., and Lee, M.-Y. 2019. Integrating Management Strategy Evaluation into fisheries management: advancing best practices for stakeholder inclusion based on an MSE for Northeast US Atlantic herring. *Can. J. Fish. Aquat. Sci.* **76**. [This issue]. doi:10.1139/cjfas-2018-0125.
- Finnoff, D., and Tschirhart, J. 2003. Harvesting in an eight-species ecosystem. *J. Environ. Econ. Manage.* **45**(3): 589–611. doi:10.1016/S0095-0696(02)00025-6.
- Francis, R.I.C.C. 1992. Use of risk analysis to assess fishery management strategies: A case study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. *Can. J. Fish. Aquat. Sci.* **49**(5): 922–930. doi:10.1139/f92-102.
- Gaichas, S.K., Fogarty, M., Fay, G., Gamble, R., Lucey, S., and Smith, L. 2017. Combining stock, multispecies, and ecosystem level fishery objectives within an operational management procedure: simulations to start the conversation. *ICES J. Mar. Sci.* **74**(2): 552–565. doi:10.1093/icesjms/fsw119.
- Golet, W.J., Galuardi, B., Cooper, A.B., and Lutcavage, M.E. 2013. Changes in the distribution of Atlantic Bluefin tuna (*Thunnus thynnus*) in the Gulf of Maine 1979–2005. *PLoS ONE*, **8**(9): e75480. doi:10.1371/journal.pone.0075480. PMID:24069420.
- Golet, W., Record, N., Lehuta, S., Lutcavage, M., Galuardi, B., Cooper, A., and Pershing, A. 2015. The paradox of the pelagics: Why bluefin tuna can go hungry in a sea of plenty. *Mar. Ecol. Prog. Ser.* **527**: 181–192. doi:10.3354/meps11260.
- Hall, C.S., Kress, S.W., and Griffin, C.R. 2000. Composition, spatial and temporal variation of common and Arctic Tern chick diets in the Gulf of Maine. *Waterbirds*, **23**(3): 430. doi:10.2307/1522180.
- Haltuch, M.A., Punt, A.E., and Dorn, M.W. 2008. Evaluating alternative estimators of fishery management reference points. *Fish. Res.* **94**(3): 290–303. doi:10.1016/j.fishres.2008.01.008.
- Haltuch, M.A., Punt, A.E., and Dorn, M.W. 2009. Evaluating the estimation of fishery management reference points in a variable environment. *Fish. Res.* **100**(1): 42–56. doi:10.1016/j.fishres.2009.03.001.
- Hatch, J.J. 2002. Arctic Tern (*Sterna paradisaea*). In *The birds of North America Online*. doi:10.2173/bna.707.
- Hayes, S., Josephson, E., Maze-Foley, K., and Rosel, P.E. (Editors). 2017. US Atlantic and Gulf of Mexico Marine Mammal Stock Assessments — 2016 [online]. NOAA Technical Memorandum NMFS-NE-241. US Department of Commerce. Available from <https://repository.library.noaa.gov/view/noaa/14864> [accessed 7 June 2018].
- Hilborn, R., and Walters, C.J. 2003. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Springer Science & Business Media.
- Hilborn, R., Amoroso, R.O., Bogazzi, E., Jensen, O.P., Parma, A.M., Szuwalski, C., and Walters, C.J. 2017. When does fishing forage species affect their predators? *Fish. Res.* **191**: 211–221. doi:10.1016/j.fishres.2017.01.008.
- Holland, D.S., and Sutinen, J.G. 1999. An empirical model of fleet dynamics in New England trawl fisheries. *Can. J. Fish. Aquat. Sci.* **56**(2): 253–264. doi:10.1139/f98-169.
- Hutniczak, B. 2014. Increasing pressure on unregulated species due to changes in individual vessel quotas: an empirical application to trawler fishing in the Baltic Sea. *Mar. Resour. Econ.* **29**(3): 201–217. doi:10.1086/677760.
- ICCAT. 2015. Report of the 2014 Atlantic bluefin tuna stock assessment session (Madrid, Spain, 22–27 September 2014). Collected Volume Scientific Papers **71**(2): 692–945. Available from https://www.iccat.int/Documents/Meetings/Docs/2014_BFT_ASSESS-ENG.pdf.
- Irwin, B.J., Wilberg, M.J., Bence, J.R., and Jones, M.L. 2008. Evaluating alternative harvest policies for yellow perch in southern Lake Michigan. *Fish. Res.* **94**(3): 267–281. doi:10.1016/j.fishres.2008.05.009.
- Jin, D., Hoagland, P., and Dalton, T.M. 2003. Linking economic and ecological models for a marine ecosystem. *Ecol. Econ.* **46**(3): 367–385. doi:10.1016/j.ecolecon.2003.06.001.
- Jin, D., Hoagland, P., Dalton, T.M., and Thunberg, E.M. 2012. Development of an integrated economic and ecological framework for ecosystem-based fisheries management in New England. *Prog. Oceanogr.* **102**: 93–101. doi:10.1016/j.pocean.2012.03.007.
- Jin, D., DePiper, G., and Hoagland, P. 2016. Applying portfolio management to implement Ecosystem-Based Fishery Management (EBFM). *N. Am. J. Fish. Manage.* **36**: 652–669. doi:10.1080/02755947.2016.1146180.
- Johnston, R.J., and Rosenberger, R.S. 2010. Methods, trends and controversies in contemporary benefit transfer. *J. Econ. Surv.* **24**(3): 479–510. doi:10.1111/j.1467-6419.2009.00592.x.
- Katsukawa, T. 2004. Numerical investigation of the optimal control rule for decision-making in fisheries management. *Fish. Sci.* **70**(1): 123–131. doi:10.1111/j.1444-2906.2003.00780.x.
- Kell, L., Pilling, G., Kirkwood, G., Pastoors, M., Mesnil, B., Korsbrekke, K., Abaunza, P., Aps, R., Biseau, A., Kunzlik, P., Needle, C., Roel, B., and Ulrich, C. 2006. An evaluation of multi-annual management strategies for ICES roundfish stocks. *ICES J. Mar. Sci.* **63**: 12–24. doi:10.1016/j.icesjms.2005.09.003.
- Kirkley, J.E., Walden, J., and Färe, R. 2011. A general equilibrium model for Atlantic herring (*Clupea harengus*) with ecosystem considerations. *ICES J. Mar. Sci.* **68**(5): 860–866. doi:10.1093/icesjms/fsr023.
- Larkin, S.L., and Sylvia, G. 1999. Intrinsic fish characteristics and intraseason production efficiency: a management-level bioeconomic analysis of a commercial fishery. *Am. J. Agric. Econ.* **81**(1): 29–43. doi:10.2307/1244448.
- Larson, D.M., Shaikh, S.L., and Layton, D.F. 2004. Revealing preferences for leisure time from stated preference data. *Am. J. Agric. Econ.* **86**(2): 307–320. doi:10.1111/j.0092-5853.2004.00580.x.
- Legault, C.M., and Palmer, M.C. 2016. In what direction should the fishing mor-

- tality target change when natural mortality increases within an assessment? *Can. J. Fish. Aquat. Sci.* **73**(3): 349–357. doi:10.1139/cjfas-2015-0232.
- Lehuta, S., Holland, D.S., and Pershing, A.J. 2014. Investigating interconnected fisheries: a coupled model of the lobster and herring fisheries in New England. *Can. J. Fish. Aquat. Sci.* **71**(2): 272–289. doi:10.1139/cjfas-2013-0185.
- Lew, D.K. 2015. Willingness to pay for threatened and endangered marine species: a review of the literature and prospects for policy use. *Front. Mar. Sci.* **2**: 96. doi:10.3389/fmars.2015.00096.
- Lew, D.K., and Wallmo, K. 2017. Temporal stability of stated preferences for endangered species protection from choice experiments. *Ecol. Econ.* **131**: 87–97. doi:10.1016/j.ecolecon.2016.08.009.
- Lew, D.K., Layton, D.F., and Rowe, R.D. 2010. Valuing enhancements to endangered species protection under alternative baseline futures: The case of the steller sea lion. *Mar. Resour. Econ.* **25**(2): 133–154. doi:10.5950/0738-1360-25.2.133.
- Li, Y., Bence, J.R., and Brenden, T.O. 2016. The influence of stock assessment frequency on the achievement of fishery management objectives. *N. Am. J. Fish. Manage.* **36**(4): 793–812. doi:10.1080/02755947.2016.1167145.
- Link, J. 2002. Does food web theory work for marine ecosystems? *Mar. Ecol. Progr. Ser.* **230**: 1–9. doi:10.3354/meps230001.
- Link, J., Griswold, C., Methratta, E., and Gunnard, J. (Editors). 2006. Documentation for the Energy Modeling and Analysis eExercise (EMAX). Northeast Fish. Sci. Cent. Ref. Doc. 06-15. US Dept. Commer., National Marine Fisheries Service, Woods Hole, Mass.
- Link, J., Overholtz, W., O'Reilly, J., Green, J., Dow, D., Palka, D., Legault, C., Vitaliano, J., Guida, V., Fogarty, M., Brodziak, J., Methratta, L., Stockhausen, W., Col, L., and Griswold, C. 2008. The Northeast U.S. continental shelf Energy Modeling and Analysis exercise (EMAX): ecological network model development and basic ecosystem metrics. *J. Mar. Syst.* **74**(1–2): 453–474. doi:10.1016/j.jmarsys.2008.03.007.
- Link, J., Col, L., Guida, V., Dow, D., O'Reilly, J., Green, J., Overholtz, W., Palka, D., Legault, C., Vitaliano, J., Griswold, C., Fogarty, M., and Friedland, K. 2009. Response of balanced network models to large-scale perturbation: implications for evaluating the role of small pelagics in the Gulf of Maine. *Ecol. Modell.* **220**(3): 351–369. doi:10.1016/j.ecolmodel.2008.10.009.
- Logan, J.M., Golet, W.J., and Lutcavage, M.E. 2015. Diet and condition of Atlantic bluefin tuna (*Thunnus thynnus*) in the Gulf of Maine, 2004–2008. *Environ. Biol. Fishes.* **98**(5): 1411–1430. doi:10.1007/s10641-014-0368-y.
- Loomis, J.B., and White, D.S. 1996. Economic benefits of rare and endangered species: summary and meta-analysis. *Ecol. Econ.* **18**(3): 197–206. doi:10.1016/0921-8009(96)00029-8.
- Lovvorn, J.R., De La Cruz, S.E.W., Takekawa, J.Y., Shaskey, L.E., and Richman, S.E. 2013. Niche overlap, threshold food densities, and limits to prey depletion for a diving duck assemblage in an estuarine bay. *Mar. Ecol. Progr. Ser.* **476**: 251–268. doi:10.3354/meps10104.
- McCConnell, K.E., and Strand, I.E. 2000. Hedonic Prices for Fish: Tuna Prices in Hawaii. *Am. J. Agric. Econ.* **82**(1): 133–144. doi:10.1111/0002-9092.00011.
- Navrud, S., and Ready, R.C. 2007. Environmental value transfer: issues and methods. Springer.
- Nisbet, I.C.T. 2002. Common Tern (*Sterna hirundo*). The birds of North America Online. doi:10.2173/bna.618.
- Northeast Fisheries Science Center. 2012. 54th Northeast Regional Stock Assessment Workshop (54th SAW) Assessment Report. US Dept of Commerce, Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543, USA.
- Overholtz, W.J., and Link, J.S. 2007. Consumption impacts by marine mammals, fish, and seabirds on the Gulf of Maine–Georges Bank Atlantic herring (*Clupea harengus*) complex during the years 1977–2002. *ICES J. Mar. Sci.* **64**(1): 83–96. doi:10.1093/icesjms/fsi009.
- Overholtz, W.J., Link, J.S., and Suslowicz, L.E. 2000. Consumption of important pelagic fish and squid by predatory fish in the northeastern U.S.A. shelf ecosystem with some fishery comparisons. *ICES J. Mar. Sci.* **57**(4): 1147–1159. doi:10.1006/jmsc.2000.0802.
- Pesaran, M.H., Shin, Y., and Smith, R.J. 2001. Bounds testing approaches to the analysis of level relationships. *J. Appl. Econ.* **16**(3): 289–326. doi:10.1002/jae.616.
- Pikitch, E., Boersma, P., Boyd, I., Conover, D., Cury, P., Essington, T., Heppell, S., Houde, E., Mangel, M., Pauly, D., Plagányi, É., Sainsbury, K., and Steneck, R. 2012. Little fish, big impact: managing a crucial link in ocean food webs. Lenfest Oceans Program, Washington, D.C.
- Plagányi, É.E., and Butterworth, D.S. 2012. The Scotia Sea krill fishery and its possible impacts on dependent predators: modeling localized depletion of prey. *Ecol. Appl.* **22**(3): 748–761. doi:10.1890/11-0441.1. PMID:22645808.
- Plagányi, É.E., Punt, A.E., Hillary, R., Morello, E.B., Thébaud, O., Hutton, T., Pillans, R.D., Thorson, J.T., Fulton, E.A., Smith, A.D.M., Smith, F., Bayliss, P., Haywood, M., Lyne, V., and Rothlisberg, P.C. 2014. Multispecies fisheries management and conservation: tactical applications using models of intermediate complexity. *Fish. Fish.* **15**(1): 1–22. doi:10.1111/j.1467-2979.2012.00488.x.
- Porch, C.E., and Laretta, M.V. 2016. On making statistical inferences regarding the relationship between spawners and recruits and the irresolute case of western Atlantic bluefin tuna (*Thunnus thynnus*). *PLoS ONE*, **11**(6): e0156767. doi:10.1371/journal.pone.0156767. PMID:27272215.
- Punt, A.E., Smith, A.D., and Cui, G. 2002. Evaluation of management tools for Australia's south east fishery. 3. Towards selecting appropriate harvest strategies. *Mar. Freshw. Res.* **53**(3): 645–660. doi:10.1071/MF01009.
- Punt, A.E., Dorn, M.W., and Haltuch, M.A. 2008. Evaluation of threshold management strategies for groundfish off the U.S. West Coast. *Fish. Res.* **94**(3): 251–266. doi:10.1016/j.fishres.2007.12.008.
- Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A., and Haddon, M. 2016a. Management strategy evaluation: Best practices. *Fish. Fish.* **17**(2): 303–334. doi:10.1111/faf.12104.
- Punt, A.E., MacCall, A.D., Essington, T.E., Francis, T.B., Hurtado-Ferro, F., Johnson, K.F., Kaplan, I.C., Koehn, L.E., Levin, P.S., and Sydemann, W.J. 2016b. Exploring the implications of the harvest control rule for Pacific sardine, accounting for predator dynamics: a MICE model. *Ecol. Modell.* **337**: 79–95. doi:10.1016/j.ecolmodel.2016.06.004.
- R Core Team. 2016. R: a language and environment for statistical computing [online]. R Foundation for Statistical Computing, Vienna, Austria. Available from <https://www.R-project.org/>.
- Rago, P., and Sosebee, K. 2010. Biological reference points for spiny dogfish [online]. Northeast Fish. Sci. Cent. Ref. Doc. 10-06: 52. Available from <https://www.nefsc.noaa.gov/publications/crd/crd1006/crd1006.pdf> [accessed 12 February 2019].
- Rago, P., and Sosebee, K. 2015. Update on the status of spiny dogfish in 2013 and projected harvests at the Fmsy proxy and Pstar of 40% [online]. Mid Atlantic Fishery Management Council Scientific and Statistical Committee. Available from https://www.mafmc.org/s/2015-Status-Report-and-Projections_final.pdf [accessed 15 November 2016].
- Rago, P.J., Sosebee, K.A., Brodziak, J.K.T., Murawski, S.A., and Anderson, E.D. 1998. Implications of recent increases in catches on the dynamics of Northwest Atlantic spiny dogfish (*Squalus acanthias*). *Fish. Res.* **39**(2): 165–181. doi:10.1016/S0165-7836(98)00181-7.
- Reimer, M.N., Abbott, J.K., and Wilen, J.E. 2017. Fisheries production: Management institutions, spatial choice, and the quest for policy invariance. *Mar. Resour. Econ.* **32**(2): 143–168. doi:10.1086/690678.
- Restrepo, V.R., Diaz, G.A., Walter, J.F., Neilson, J.D., Campana, S.E., Secor, D., and Wingate, R.L. 2010. Updated estimate of the growth curve of Western Atlantic bluefin tuna. *Aquat. Living Resour.* **23**(4): 335–342. doi:10.1051/alr/2011004.
- Richardson, L., and Loomis, J. 2009. The total economic value of threatened, endangered and rare species: an updated meta-analysis. *Ecol. Econ.* **68**(5): 1535–1548. doi:10.1016/j.ecolecon.2008.10.016.
- Scopel, L., and Diamond, A. 2017a. Predation and food–weather interactions drive colony collapse in a managed metapopulation of Arctic Terns (*Sterna paradisaea*). *Can. J. Zool.* **96**(1): 13–22. doi:10.1139/cjz-2016-0281.
- Scopel, L.C., and Diamond, A.W. 2017b. The case for lethal control of gulls on seabird colonies. *J. Wildl. Manage.* **81**(4): 572–580. doi:10.1002/jwmg.21233.
- Scopel, L.C., Diamond, A.W., Kress, S.W., Hards, A.R., and Shannon, P. 2018. Seabird diets as bioindicators of Atlantic herring recruitment and stock size: a new tool for ecosystem-based fisheries management. *Can. J. Fish. Aquat. Sci.* **75**(8): 1215–1229. doi:10.1139/cjfas-2017-0140.
- Smith, B.E., and Link, J.S. 2010. The trophic dynamics of 50 finfish and 2 squid species on the Northeast US Continental Shelf [online]. NOAA Technical Memorandum NMFS-NE-216. National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, USA. Available from <http://www.nefsc.noaa.gov/publications/tm/tm216/> [accessed 26 April 2016].
- Smith, L.A., Link, J.S., Cadrin, S.X., and Palka, D.L. 2015. Consumption by marine mammals on the Northeast U.S. continental shelf. *Ecol. Appl.* **25**(2): 373–389. doi:10.1890/13-1656.1. PMID:26263661.
- Stouffer, S.A., Suchman, E.A., DeViney, L.C., Star, S.A., and Williams, R.M., Jr. 1949. The American soldier: adjustment during army life. (Studies in social psychology in World War II, Vol. 1.) Princeton University Press.
- Sydemann, W.J., Thompson, S.A., Santora, J.A., Koslow, J.A., Goericke, R., and Ohman, M.D. 2015. Climate–ecosystem change off southern California: time-dependent seabird predator–prey numerical responses. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **112**: 158–170. doi:10.1016/j.dsr2.2014.03.008.
- Sydemann, W.J., Thompson, S.A., Anker-Nilssen, T., Arimitsu, M., Bennison, A., Bertrand, S., Boersch-Supan, P., Boyd, C., Bransome, N.C., Crawford, R.J., Daunt, F., Furness, R.W., Gianuca, D., Gladics, A., Koehn, L., Lang, J.W., Løgerwell, E., Morris, T.L., Phillips, E.M., Provencher, J., Punt, A.E., Sarau, C., Shannon, L., Sherley, R.B., Simeone, A., Wanless, R.M., Wanless, S., and Zador, S. 2017. Best practices for assessing forage fish fisheries–seabird resource competition. *Fish. Res.* **194**: 209–221. doi:10.1016/j.fishres.2017.05.018.
- Tschirhart, J. 2000. General equilibrium of an ecosystem. *J. Theor. Biol.* **203**(1): 13–32. doi:10.1006/jtbi.1999.1058. PMID:10677274.
- Whitlock, M.C. 2005. Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *J. Evol. Biol.* **18**(5): 1368–1373. doi:10.1111/j.1420-9101.2005.00917.x. PMID:16135132.
- Wiedenmann, J., Wilberg, M., Sylvia, A., and Miller, T. 2017. An evaluation of acceptable biological catch harvest control rules designed to limit overfishing. *Can. J. Fish. Aquat. Sci.* **74**(7): 1028–1040. doi:10.1139/cjfas-2016-0381.
- Zador, S.G., Gaichas, S.K., Kasperski, S., Ward, C.L., Blake, R.E., Ban, N.C., Himes-Cornell, A., and Koehn, J.Z. 2017. Linking ecosystem processes to communities of practice through commercially fished species in the Gulf of Alaska. *ICES J. Mar. Sci.* **74**(7): 2024–2033. doi:10.1093/icesjms/fsx054.