Seabird diets as bioindicators of Atlantic herring recruitment and stock size: a new tool for ecosystem-based fisheries management

Lauren C. Scopel, Antony W. Diamond, Stephen W. Kress, Adrian R. Hards, and Paula Shannon

Abstract: Ecosystem-based fishery management requires understanding of relationships between exploited fish and their predators, such as seabirds. We used exploratory regression analyses to model relationships between Atlantic herring (Clupea harengus) in the diet of seabird chicks at nine nesting colonies in the Gulf of Maine and four types of fishery- and survey-derived herring data. We found several strong relationships, which suggests spatial structuring in herring stocks and likely patterns of herring movements before they recruit into the fishery. Some types of herring data seldom used in stock assessments — notably acoustic surveys, fixed-gear landings, and mass-at-age (i.e., weight-at-age) — correlated as strongly with seabird data as more commonly used series, such as mobile-gear landings and modeled spawning stock biomass. Seabird chick diets collected at specific locations thus offer a promising means to assess the size, distribution, and abundance of juvenile herring across a broad area prior to recruitment, which is a major source of uncertainty in fisheries. Common terms (Sterna hirundo) showed the most potential as a bioindicator, correlating well and showing consistent spatial patterns with 11 of 13 fishery data series.

Introduction

Fisheries managers worldwide are increasingly expected to oversee harvested stocks with due regard to their interactions with other species in the ecosystem (Stephenson 2012). Successful application of this approach, often referred to as ecosystem-based fishery management or EBFM, is widely recognised as essential to the maintenance of healthy marine ecosystems and the communities they support (Christensen et al. 1996; Dickey-Collas et al. 2013; Skern-Mauritzen et al. 2016).

A fundamental feature of EBFM is understanding the multiple interactions among constituent species (Travis et al. 2014), including (but not limited to) harvested species and their prey. Fishery harvest is treated as a key ecosystem service to be sustained, yet many other predators rely on the same prey species, and sustaining their populations is an equally valid ecosystem service (Diamond 2012; Fogarty 2014). The idea that the needs of natural predators “deserve” consideration alongside those of fisheries is increasingly becoming accepted (Travis et al. 2014) and incorporated into both global biodiversity goals (Convention on Biological Diversity [No date]) and national fishery-management policies (Stephenson 2012; Skern-Mauritzen et al. 2016).

Seabirds are among the most widespread and best-studied predators of marine fish, and they are affected by large-scale changes in abundance of prey stocks especially as a result of overfishing (Duffy 1983; Cury et al. 2011; Cook et al. 2014). Seabirds may also reflect other changes in the marine environment, including contamination and oceanographic and climatic change, leading to well-established use of “seabirds as environmental indicators” or SEI (Cairns 1987; Diamond and Devlin 2003; Einoeder 2009). Most of the SEI literature examines various behavioral responses — such as...
time budgets (Jodice et al. 2006) or reproductive success (Monaghan et al. 1989; Wanless et al. 2005) — to changing food availability, but Parsons et al. (2008) pointed out that seabird chick diet data offer “the most immediate signal of availability”, which are not widely available because they are time-consuming to collect. We take advantage of our long time series of seabird chick diet data to conduct exploratory analyses using data routinely collected to manage one widely harvested forage fish, Atlantic herring (Clupea harengus), which has been prominent in the diet of seabirds and many other top predators (see below).

Forage fish are the target of many fisheries throughout the world (Pikitch et al. 2012). They are generally small or medium-sized, pelagic, schooling species, often with high fat content (Harris and Hislop 1978; Hislop et al. 1991; Wanless et al. 2005), characterized ecologically by their position midway in the food chain between phytoplankton and larger predators; in such marine systems the energy required by higher trophic levels is channeled through the forage fish (Pikitch et al. 2012). In such “wasp-waist” systems (Bakun et al. 2009), the forage fish play a “keystone” role (i.e., a species that exerts influences out of proportion to its abundance; sensu Paine 1995; Johnson et al. 2017), supporting a variety of top predators, including larger fish, marine mammals, and marine birds; if these forage fish are the target of a fishery, it is critically important to avoid overfishing because of all the predators that rely on the same prey as the fishery (Curry et al. 2011). If such forage fish show high caloric density (i.e., are high in fat content), they are particularly likely to be targeted by optimal foragers and function as keystone prey.

In systems where seabirds feed on forage fish that are also the target of a fishery, there is enhanced potential both for negative impacts of the fishery on the birds and for the birds to respond quickly to changes in the fish stock; much of the SEI literature refers to such systems. Among many examples are sandeels (or Ammodytes spp.) in the North Sea (Monaghan et al. 1989; Frederiksen et al. 2008; Cook et al. 2014) and Gulf of Alaska (Jodice et al. 2006; Suryan et al. 2006); capelin (Mallotus villosus) around Newfoundland and Labrador (Montevecchi and Myers 1996; Regehr and Montevecchi 1997), the Barents Sea (Barrett and Furness 1990), and Gulf of Alaska (Jodice et al. 2006); anchovy and sardines in the Benguela Current (Crawford et al. 2008; Green et al. 2015), the Humboldt Current (Duffy 1983), and California Current (Mills et al. 2007; Sydeman et al. 2009; Thompson et al. 2012); and Atlantic herring in Iceland, Norway (Anker-Nilssen 1987), and the Gulf of Maine (Diamond and Devlin 2003; Breton and Diamond 2014).

Atlantic herring in the Gulf of Maine play a widely recognised keystone role (Kenney et al. 1997; Bakun et al. 2009; Diamond 2012) as a forage fish. Predators with a predominantly herring diet include bluefin tuna (Thunnus thynnus), themselves the object of an important fishery (Golet et al. 2007, 2015), humpback (Megaptera novaeangliae) and fin (Balaenoptera physalus) whales, other cetaceans, seals (Kenney et al. 1997), and several species of long-distance migratory seabirds as well as those breeding locally (Diamond 2012). However, different predators consume different sizes (and therefore age classes) of the herring stock. Juvenile herring hatched in fall are referred to as “0-group” until 1 January when they become 1-group (or 1-year-olds). Herring reach sexual maturity at 3–4 years but can enter the fishery at age-2 (Stephenson et al. 1993; Fig. 1). Herring in the Gulf of Maine spawn mostly in late summer or fall, so those eaten by seabirds breeding the following summer are most often 1 year old, though the larger auks (razorbill (Alca torda) and common murre (Uria aalge)) also take 2-year-old fish. Seabirds sample herring younger than 2 years old, offering the possibility that data on their diets could predict recruitment and aid in stock assessment and EBFM.

Our focus in this paper is on assessing possible relationships between the herring content in the diet of seabird chicks and various measures of herring abundance, with a view to identifying possible predictors of future herring stock size. We assume that seabirds take preferred prey (in this case, 1-year-old herring) in relation to the abundance of that prey; thus, a high proportion of herring in the diet reflects an abundance of herring within the foraging range of the colony. Specifically, we assess relationships between sets of herring data commonly used in stock assessments in both Canadian and American parts of the Gulf of Maine and several measures of diets of breeding seabirds and their chicks. From this assessment, we examine the potential of diet data from seabird species and colonies as indicators of value to future EBFM. We include diet data from multiple species at widely separated colonies, allowing us to address both the different spatial and temporal scales at which seabirds and fisheries operate and the varied foraging behaviours of different species of seabird (Greenstreet et al. 2000; Becker and Chapdelaine 2003).
analyses are driven by two broad hypotheses: (i) the diets of seabirds indicate the strength of a herring cohort at age-1, which predicts future stock size (seabird diets predict fishery), and (ii) data on breeding herring from fisheries predict future productivity and the prevalence of herring in future seabird diets (fishery predicts seabird diets). To reflect these hypotheses and the temporal disjuncture between the age-1 cohorts targeted by seabirds and the older ages targeted by various parts of the fishery, we perform an exploratory analysis examining relationships between fishery and seabird data at lags ranging from −3 (fishery predicts seabirds 3 years later) to +5 (seabirds predict fishery 5 years later).

We identify the strongest seabird–fishery relationships, rank herring data for their relevance to herring predators, and identify seabird colonies and species as potentially useful indicators of the herring stock.

Methods

Seabird study area and diet data collection

Seabird diet data were collected from nine colonies in the Gulf of Maine and Bay of Fundy region, where long-term seabird monitoring programs have been conducted since the early 1970s (Fig. 2). Seven Maine colonies (Stratton Island, Outer Green Island, Jenny Island, Pond Island National Wildlife Refuge (NWR), Eastern Egg Rock, Matinicus Rock, and Seal Island NWR) are managed by National Audubon Society’s Seabird Restoration Program (SRP), one is managed by the US Fish and Wildlife Service (USFWS; Petit Manan Island, Maine), and one (Machias Seal Island (MSI), New Brunswick) is managed by the Canadian Wildlife Service and monitored by the University of New Brunswick (UNB). Protocols to collect diet data were similar across islands. The Animal Care
Committee at UNB reviewed annual procedures performed at MSI, and all SRP and USFWS data collection was in accordance with federal banding permits. We consider four colonies as “western” (Stratton Island, Outer Green Island, Jenny Island, and Pond Island NWR), three colonies as “central” (Eastern Egg Rock, Matinicus Rock, and Seal Island NWR), and two colonies as “eastern” (Petit Manan Island, MSI).

We included diet data from four species: two surface-feeding tern species (common tern (Sterna hirundo) and Arctic tern (Sterna paradisaea)) and two diving alcid species (Atlantic puffin (Fratercula arctica) and razorbill (Alca torda)). Common terns breed on all nine colonies, and we used data from Arctic terns on four (Matinicus Rock, Seal Island, Petit Manan Island, and MSI). Atlantic puffins on four (Eastern Egg Rock, Matinicus Rock, Seal Island, and MSI), and razorbills on two (Matinicus Rock and MSI). The latter three species tended to nest in colonies in the central and eastern parts of the Gulf of Maine, representing the southern limit of their North American ranges.

Feeding watches were performed throughout the chick-rearing period of each study species on each island, from multiple observation blinds erected in each colony. Observers used binoculars to identify common prey species at each colony, estimating prey length relative to the adult’s bill length. For each feeding, observers recorded the nest number, prey recipient, number of prey, prey length, and prey species. Further details on seabird diet observations are included in the online Supplementary Data S1.

A full summary of the time series for all species at all colonies is included in Table S1. Time series at SRP colonies begin in 1991 or 1992 for Arctic terns, 2002 or 2003 for common terns, and 2005 for alcids. Time series at Petit Manan Island began in 2000 for Arctic terns and 2002 for common terns. At MSI, data begin in 1990 for Arctic terns and 1995 for common terns and alcids. The tern colony at MSI collapsed in 2006 (Diamond 2009; Gaston et al. 2009), so no tern feeding observations could be performed from 2006 to 2013. The diet study for Arctic terns resumed in 2014 with the restoration of the colony (Scopel and Diamond 2017), and diet estimates for common terns were derived from video observations in 2016.

Seabird diet data treatment

Seabird diet data treatment follows that performed in Scopel and Diamond (2018). All raw feeding data were converted into annual estimates of wet mass. For each species at each colony, taxa that made up at least 1% of the diet by number were retained for conversion into mass (Table S2). If possible, length–mass conversion formulae were developed from prey samples collected opportunistically by UNB at MSI as part of their annual protocol (available from http://www.unb.ca/research/alar/msi-seabirds/protocol-methods.html); prey samples were measured for total length to the nearest 0.1 mm and wet mass to the nearest 0.1 g. For prey taxa with too few samples to develop conversions (n = 8), we used values from the literature (Table S3). Once all prey conversions were developed, we applied the conversion formulae to our annual diet data for each colony–species pair (e.g., common tern – Jenny Island, Atlantic puffin – Seal Island) to derive estimates of wet mass for each prey taxon and for each year. Sums of mass for each prey taxon were tallied and summarized for each year.

For this study, we calculated four representations of herring in the seabird diet: two rates (herring mass per feeding, herring mass per observed chick-hour; i.e., the total number of hours each chick was observed at a nest (tern) or plot-hour (alcid)), proportion of herring in the diet, and mean mass of individual herring. A summary of these representations of seabird diet, including sample sizes, can be found in Tables S4–S12 in the Supplementary Data. Strength of herring cohort by colony–species pair is included in Fig. 3.

Herring stock data sources

Herring in the Gulf of Maine and Bay of Fundy region are managed by an American and a Canadian fishery, respectively; although both countries manage their quotas and stocks independently, they do share data and issue status reports jointly through the Transboundary Resource Assessment Committee (TRAC). The Canadian fishery manages the Bay of Fundy and Scotian Shelf herring as the 4WX herring stock complex, requiring the separation of Bay of Fundy herring (4X) from reported totals in the 4VWX reports. The American fishery identifies the Gulf of Maine and Georges Bank as separate spawning components (5Y and 5Z, respectively), but both are modeled as part of the same stock complex (NEFSC 2012).

Owing to the diversity of data that can represent a fish stock, we considered a variety of potential herring data types for this analysis. We selected four broad categories that represent different aspects of the herring fishery: fishery-independent acoustic surveys of spawning stock biomass, landings, weights-at-age (herein referred to as mass-at-age), and output from stock assessment models. Although landings are relatively simple to measure and data are abundant, they do represent targeted and technologically assisted sampling of fish that does not necessarily represent the behaviour of small piscivores, such as seabirds. We therefore included data from acoustic surveys to obtain a fishery-independent estimate of spawning stock biomass. The Canadian fishery has reported declining average masses-at-age since the 1970s (Fisheries and Oceans Canada (DFO) 2015), so we included Canadian mass data to reflect this potentially important property of the stock. Finally, we included two types of output from stock assessment models. Age-1 recruitment was estimated by both the American and Canadian stock assessments via forward-cast (age-structured assessment program with a Beverton–Holt stock–recruitment relationship) and back-cast (virtual population assessment, no longer used officially) age-structured models, respectively; these serve as the best available data to estimate the abundance of 1-year-old herring in the fishery, which can be compared directly with seabird diets. Finally, we included estimates of spawning stock biomass from the American stock assessment model for comparison with landings and acoustic surveys.

Landings data were obtained from stock assessment reports. We included mobile-gear landings (primarily purse seine and midwater trawl) from the American fishery (Deroba 2015), mobile-gear landings from New Brunswick (Stephenson et al. 1998; Singh et al. 2016), combined landings (American mobile-gear landings and all fixed-gear landings, Deroba 2015), and fixed-gear landings from southwestern New Brunswick (Deroba 2015). Most fixed-gear landings in the region are caught in weirs in southwestern New Brunswick.

Acoustic data were obtained for the Canadian and American stocks; Canadian data were obtained from stock assessment reports (DFO 2015), and American data were obtained directly from M. Jech at the NEFSC, following the format published in the American stock assessment report (NEFSC 2012). Canadian surveys sampled herring at Trinity Ledge, German Bank, and Scots Bay, Nova Scotia; American surveys sampled herring at Georges Bank (Fig. 2). Acoustic survey data in both regions began in 1999, and continued through 2012 in the USA and 2014 in Canada.

Mass-at-age data were obtained directly from M. Power at DFO. We selected three types of mass data: mass-at-age estimates for age-1 fish (the youngest available, the same age that most of the seabirds would consume), fish aged 3–5 years (which represent the majority of the spawning stock), and fish aged 3+ years. The
majority of spawners in the stock are 3–5 years old, but the entire spawning component can reach age-10+. We compared the importance of just young spawners (3–5 years) and the entire spawning stock; although younger spawners are more abundant, older and larger fish may contribute more to future recruitment owing to greater fecundity (Wootton 1998; Hsieh et al. 2006). DFO also identified a substantial decline in the average growth of herring between ages 2 and 3 in recent years (M. Power, DFO, St. Andrews, New Brunswick, personal communication, 2016), which we included as a fourth data type in our mass category as “growth increment.”

Finally, we included spawning stock biomass estimates from the American model (Deroba 2015) and estimates of the number of age-1 herring as predicted by the American (NEFSC 2012) and Canadian (obtained directly from R. Singh, DFO) models.

Complete herring data used in this analysis, including years of available data, are included in Tables S13–S14 in the Supplemental Data. Strength of herring cohorts, as measured by the American stock assessment model, are included in Fig. 3.

Data analysis — exploratory assessment of lags

Herring sampled by seabirds and by the fishery are rarely of the same age class; most seabirds consume 1- or 2-year-old herring, but the herring fishery targets fish of ages 2+ (fixed gear) and 3+ years old (mobile gear, acoustic surveys; Fig. 1). This age discrepancy offers the potential to predict future stock data from the younger age classes consumed by seabirds. To investigate potentially useful predictive lags, we identified significant lags using the cross-correlation function (ccf(), package “stats”) in R (version 3.1.2; R Core Team 2014). We examined all lags between −3 to +5 for each seabird colony and fishery data pair; positive lags indicate that seabird diets predict fishery data, and vice versa for negative lags. Any lag that was deemed significant by the cross-correlation function (>0.4; R Core Team 2014) was retained for further analysis. We identified three classes of lags: (i) no lag or lags of +1 years (“no lags” or “short lags”), (ii) lags of +2 to +5 years (long lags), and (iii) lags of −1 to −3 years (negative lags).

We used 19 colony–species pairs of seabird diet data (e.g., common tern on Matinicus Rock, Atlantic puffin at Eastern Egg Rock), each with four representations of seabird-derived herring (e.g., herring mass per feeding), and correlated them to 13 variables derived from the herring fishery (e.g., fixed-gear landings). For mean mass of individual herring (seabird), we correlated these to the four fishery-derived mass variables only.
Data analysis — validation of significant correlations

Following the exploratory assessment of lags, we examined all significant seabird–fishery correlations in greater detail. All analyses were performed in R. All data were normalized (mean of 0, standard deviation ±1). We used univariate lagged regression to predict fishery-derived estimates of herring from seabird-derived estimates of herring. Owing to the direct comparison of two herring measures, we tested only simple linear relationships; we assumed that differences in cohort size would persist linearly throughout the time series.

Since seabirds can consume 1- or 2-year-old herring, we tested the possibility of temporal autocorrelation by including an autoregressive (AR) correlation structure within our models (Zuur et al. 2009). We prepared two sets of models for each significant lag, one with and one without an AR1 correlation structure. We used generalized least squares regression (GLS, gls()), package “nlme”; Pinheiro et al. 2016) and compared each model pair using AICc (package “AICcmodavg”, Mazerolle 2015); for each model, the structure with the lowest AICc was selected.

Models were built with fishery data as the response variable and seabird data as the predictor. Series of models were produced for each fishery variable and ranked using AICc. Any models that ranked lower than a model using only year as a predictor (i.e., the “null” model) were considered uninformative and removed from further consideration.

After fitting GLS models, we checked models for meaningful effect sizes and explained variation, using regression slopes and McFadden’s pseudo-$R^2$ (McFadden 1974), respectively. Models that scored lower than 0.5 in either category were considered uninformative and removed from further consideration. In some cases, a single seabird–fishery lag was represented multiple times with different rate and proportional data; if one model was ranked more than 3 $\Delta$AICc, units higher than another model within the same seabird–fishery lag (e.g., Seal Island puffins — acoustic +3), we removed it from consideration. All instances with lags within 3 $\Delta$AICc units were kept.

For all remaining models, leave-one-out cross-validation (Efron and Tibshirani 1993) was performed to assess the model’s predictive ability and stability of slope. Models where slopes dropped below 0.25 were considered uninformative and removed from further consideration. Once cross-validation was performed, we reduced the model set to a single instance of each seabird–fishery pair, if possible, based on prediction error (CV), pseudo-$R^2$, slope, and potential outlier years. Multiple lags for one seabird–fishery pair were retained if they were well-supported. Models with greater pseudo-$R^2$, greater slopes, and lower prediction error were considered stronger.

Relevance rankings

GLS models were separated into potential quantitatively and qualitatively useful lags; CVs of <0.5 were considered suitable for quantitative prediction, but those >0.5 were considered useful to represent general qualitative trends only. CVs of <0.2 were considered the strongest predictive models, and these were considered potentially reliable bioindicators.

Using the cross-validated GLS models, fishery data were ranked based on their potential relevance to future EFPM. Ranking criteria for each fishery data type included the number of seabird species, seabird colonies, and geographical subregions (western, central, or eastern) that were represented by cross-validated models, the number of models that passed cross-validation, and the number of models with CVs < 0.2. Model sets that were representative of a larger spatial area, highly correlated with more species, and had strong predictive value were considered the most relevant to future study.

Post hoc analyses

Following our initial analyses, we performed a post hoc comparison between our seabird variables and American mass-at-age data (NEFSC 2012). We found a marked reduction in mass-at-age after age-3 in both Canadian and US data (Fig. 4), which encouraged us to include both data sets, using the same age categories (ages 1, 3–5, and 3+) as the Canadian mass data. The American time series included data through 2011.

We also examined the importance of time series length within our longest seabird time series. We divided four time series (MSI Atlantic puffins and razorbills, Matinicus Rock and Seal Island NWR Arctic terns) into two subsets: an “early” time series including years 2005 and earlier (to emulate the length of the MSI tern series) and a “late” time series including years 2002 and later (to emulate the length of the SRP common tern series). We examined cross-correlations between these subsets and compared them with cross-correlations of the full time series to determine if changes in the seabird–fishery relationship over time weakened the strength of the observed correlations in these data. This division coincided with the timing of a regime shift between 2000 and 2004 (Smith et al. 2012) with potential consequences for nutrient status and the plankton community (Townsend et al. 2010), which can be expected to affect responses in species higher in the food web such as forage fish and their predators.

Results

Assessment of lags and validation of models

We performed 6555 seabird–fishery cross-correlation comparisons, identifying 472 seabird–fishery data pairs with significant lags during the first stage of the analysis; 71 included an autoregressive correlation structure. We performed cross-validation on 347 GLS models, retaining 70 (Table S15). Of these, 19 had CVs > 0.5, 38 had CVs between 0.2 and 0.5, and 13 had CVs < 0.2. Models with CVs < 0.2 were considered to be the highest quality (Table 1) and are examined in full detail below. Models with CVs between 0.2 and 0.5 had quantitative value but showed higher prediction error, and those with CVs > 0.5 were considered qualitatively useful only; general trends from these two lower ranked categories are also discussed below.

Top-ranked models: CV < 0.2

A complete list of the 13 top-ranked models, including their CVs, pseudo-$R^2$, and slopes, can be found in Table 1. Depictions of these models can also be found in Figs. 5 and S1–S2 in the Supplemental Data. Eight of 13 fishery data types and 10 of 19 seabird–colony pairs were involved in this model set, including all four seabird species and six of nine seabird colonies, encompassing all three geographic subregions. Acoustic surveys, landings, and model output were equally represented in this model set, but only one model included mass data. Common tern models were the most abundant, featured nine times, and six colonies of common terns, encompassing all three geographic subregions. Acoustic surveys, landings, and model output were equally represented in this model set, but only one model included mass data. Common tern models were the most abundant, featured nine times, and six colonies of common terns, encompassing all three geographic subregions. Acoustic surveys, landings, and model output were similarly represented in this model set, but only one model included mass data. Common tern models were the most abundant, featured nine times, and six colonies of common terns, encompassing all three geographic subregions. Acoustic surveys, landings, and model output were equally represented in this model set, but only one model included mass data. Common tern models were the most abundant, featured nine times, and six colonies of common terns, encompassing all three geographic subregions. Acoustic surveys, landings, and model output were equally represented in this model set, but only one model included mass data. Common tern models were the most abundant, featured nine times, and six colonies of common terns, encompassing all three geographic subregions. Acoustic surveys, landings, and model output were equally represented in this model set, but only one model included mass data.
birds. Of the remaining seabird–colony pairs, common terns were part of 25 models, puffins in 17, Arctic terns in nine, and razorbills in six.

Negative lags featured in 17 models, particularly those including mass, fixed-, or mobile-gear landings data. Long lags were in 12 models, including acoustic surveys, mobile-gear landings, spawning stock biomass, and mass data. The remaining 28 models included short or no lag and incorporated 11 different fishery variables.

### By fishery data type: all models

Acoustic surveys from the Bay of Fundy showed long lags with western common tern and central alcid data and short or no lags with terns at MSI and razorbills at Matinicus Rock. Acoustic surveys from Georges Bank linked exclusively to terns at MSI (Fig. 5) at long and short or no lags.

Fixed-gear landings had a consistent lag of −2 for four of six cross-validated models; three seabird species from MSI, both tern species from Petit Manan Island, and common terns from Matinicus Rock showed links to fixed-gear landings, which primarily occur in waters near MSI (Fig. S11). American mobile-gear landings correlated with data from seven seabird–colony pairs, spread across western and central colonies, but neither eastern nor Arctic tern data were included in these models. Lags for mobile-gear models included three long lags, two short lags, and two negative lags.

American mobile-gear landings had the second-highest number per model type, following acoustic surveys from Georges Bank in Fig. 4. Mean masses-at-age for the (A) American stock assessment and (B) Canadian stock assessment. American data include 1965–2011, while the Canadian data include 1965–2015.

### Table 1. List of seabird–fishery predictive models with a prediction error < 0.2.

<table>
<thead>
<tr>
<th>Fishery category</th>
<th>Fishery description</th>
<th>Seabird colony</th>
<th>Seabird species</th>
<th>Seabird data type</th>
<th>prediction error</th>
<th>SE</th>
<th>pseudo-R²</th>
<th>slope</th>
<th>SE</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acoustic</td>
<td>Bay of Fundy</td>
<td>Machias Seal Is.</td>
<td>Common tern</td>
<td>Proportion</td>
<td>+1</td>
<td>0.03</td>
<td>0.00</td>
<td>1.14</td>
<td>0.00</td>
<td>8</td>
</tr>
<tr>
<td>Model</td>
<td>Spawning stock</td>
<td>Matinicus Rock</td>
<td>Razorbill</td>
<td>Mass/feeding</td>
<td>+5</td>
<td>0.05</td>
<td>0.01</td>
<td>1.00</td>
<td>0.16</td>
<td>5</td>
</tr>
<tr>
<td>Landings</td>
<td>Combined</td>
<td>Machias Seal Is.</td>
<td>Common tern</td>
<td>Mass/h</td>
<td>+4</td>
<td>0.08</td>
<td>0.00</td>
<td>0.94</td>
<td>0.64</td>
<td>11</td>
</tr>
<tr>
<td>Landings</td>
<td>Fixed</td>
<td>Machias Seal Is.</td>
<td>Common tern</td>
<td>Proportion</td>
<td>−2</td>
<td>0.13</td>
<td>0.01</td>
<td>0.86</td>
<td>0.62</td>
<td>12</td>
</tr>
<tr>
<td>Model</td>
<td>Recruits, USA</td>
<td>Matinicus Rock</td>
<td>Common tern</td>
<td>Proportion</td>
<td>+5</td>
<td>0.13</td>
<td>0.01</td>
<td>0.87</td>
<td>1.07</td>
<td>9</td>
</tr>
<tr>
<td>Model</td>
<td>Spawning stock</td>
<td>Jenny Is.</td>
<td>Common tern</td>
<td>Mass/feeding</td>
<td>+5</td>
<td>0.13</td>
<td>0.01</td>
<td>0.91</td>
<td>1.07</td>
<td>8</td>
</tr>
<tr>
<td>Acoustic</td>
<td>Georges Bank</td>
<td>Machias Seal Is.</td>
<td>Common tern</td>
<td>Mass/h</td>
<td>+4</td>
<td>0.15</td>
<td>0.01</td>
<td>0.73</td>
<td>1.03</td>
<td>11</td>
</tr>
<tr>
<td>Mass</td>
<td>At age-1</td>
<td>Seal Is.</td>
<td>Puffin</td>
<td>Mean mass</td>
<td>+5</td>
<td>0.15</td>
<td>0.02</td>
<td>0.91</td>
<td>0.82</td>
<td>24</td>
</tr>
<tr>
<td>Model</td>
<td>Recruits, USA</td>
<td>Matinicus Rock</td>
<td>Puffin</td>
<td>Proportion</td>
<td>0</td>
<td>0.17</td>
<td>0.02</td>
<td>0.86</td>
<td>1.33</td>
<td>7</td>
</tr>
<tr>
<td>Acoustic</td>
<td>Bay of Fundy</td>
<td>Pond Is.</td>
<td>Common tern</td>
<td>Proportion</td>
<td>+3</td>
<td>0.18</td>
<td>0.01</td>
<td>0.76</td>
<td>0.97</td>
<td>22</td>
</tr>
<tr>
<td>Landings</td>
<td>Combined</td>
<td>Petit Manan Is.</td>
<td>Common tern</td>
<td>Proportion</td>
<td>0</td>
<td>0.18</td>
<td>0.01</td>
<td>0.80</td>
<td>0.54</td>
<td>15</td>
</tr>
<tr>
<td>Landings</td>
<td>Mobile, USA</td>
<td>Stratton Is.</td>
<td>Common tern</td>
<td>Mass/h</td>
<td>+5</td>
<td>0.18</td>
<td>0.02</td>
<td>0.87</td>
<td>0.64</td>
<td>19</td>
</tr>
<tr>
<td>Acoustic</td>
<td>Georges Bank</td>
<td>Machias Seal Is.</td>
<td>Arctic tern</td>
<td>Mass/h</td>
<td>+4</td>
<td>0.19</td>
<td>0.01</td>
<td>0.65</td>
<td>0.88</td>
<td>14</td>
</tr>
</tbody>
</table>

**Note:** Models are ranked by prediction error. All models are considered strong potential bioindicators.
of significant correlations in early model testing, but many models failed cross-validation owing to high-leverage points in 2006 and 2009.

Modeled spawning stock biomass was the most prolific fishery type in the full cross-validated model set, including 10 seabird-colony pairs at eight colonies, yet many models failed cross-validation owing to high leverage from the 2014 data point. Five models included long lags, and the remaining models had short or no lags. Recruitment data from the USA were most strongly linked to common tern and puffin data from two central colonies, but overall were much more spatially limited; all models had no lag. This indicates that recruits, as defined by the American stock assessment, are of the same age as the fish consumed by seabirds. Conversely, recruits from the Canadian stock assessment were represented primarily by +1-year lags, indicating that recruits, as defined by the Canadian stock assessment, are 1 year younger than the fish consumed by seabirds. Canadian recruitment models involved common terns, puffins, and razorbills from five central or western colonies. Although Canadian recruitment data were more widespread in their links with seabird data relative to American recruitment data, American models on average had lower prediction error; no Canadian recruitment models were included in the top-ranked model set. Common terns and puffins at Seal Island NWR were the only seabird–colony pairs to correlate with both sets of recruitment data.

Canadian mass data were included in 23 total models, but only one ranked highly and lags were inconsistent, especially in the mass-at-age-1 category. Similarly to the Canadian recruitment data, mass-at-age-1 models involved common tern, puffin, and razorbill data from central and western colonies, incorporating five colonies. The lags featured in the mass-at-age-1 data, however, were quite variable, including seven negative lags and four long lags. Mass-at-ages 3–5 data and growth increment were included in seven models, representing alcids from central colonies and terns at Petit Manan Island. Mass-at-age-3+ data were linked with terns at MSI only.

Fig. 5. Top-ranked models showing relationships between seabird diet data and American (blue) and Canadian (orange) acoustic surveys in the Gulf of Maine and Bay of Fundy region. Common and Arctic tern diet data at Machias Seal Island (MSI) predicted acoustic surveys at Georges Bank 4 years later, while common terns at Pond Island (PINWR) predicted acoustic surveys from the Bay of Fundy 3 years later. Dashed lines suggest proposed larval movements from spawning areas toward seabird colonies, where seabirds consume juvenile herring. Solid lines indicate connectivity back to natal grounds. [Colour online.]
By seabird species and colony: all models

Common terns were represented strongly; all nine colonies had models that passed cross-validation, and eight colonies were included in models with CVs < 0.5. Common tern models also involved 11 of 13 fishery data series, showing high concordance with a variety of fishery data types. The strongest models for common terns included both acoustic surveys, all three types of landings, and American recruitment and spawning stock biomass estimates. Common tern data also formed the strongest models, as ranked by CV and by pseudo-$R^2$.

Arctic tern data at MSI were linked with all four types of fishery data, including five variables, most often by +1 lags. MSI Arctic tern data correlated most frequently and most strongly with acoustic survey data, especially in Georges Bank. At Petit Manan Island, Arctic tern data were linked to just three variables, representing either mass or landings data, and relationships were weaker than the MSI models. Arctic tern data from central colonies did not correlate strongly with anything.

Atlantic puffins at all four colonies had models with CVs < 0.5, including nine fishery variables, yet only one model was included in the top-ranked set, featuring mass-at-age data. Puffins at Seal Island NWR and Matinicus Rock were the best represented in these models, showing strong relationships with mass, landings, and model data. All central colonies had models including Canadian recruitment estimates, American mobile-gear landings, and Canadian mass-at-age data. Puffins at MSI were linked only to fixed-gear landings. Razorbills at Matinicus Rock were linked to six fishery variables, including all four types of fishery data. Only one model was included in the top-ranked set, incorporating estimates of spawning stock biomass. Razorbill models were similar to puffin models from central colonies, also including Canadian recruitment estimates, mobile-gear landings, and mass-at-age data. Razorbills at MSI showed no relationship with any fishery data.

Relevance rankings

Rankings are listed in Table 2. Of the 13 fishery data series, two have high relevance to future ecosystem-based fisheries management, four have moderate relevance, six have low or limited relevance, and one has poor relevance.

Post hoc analyses

Six American mass-at-age models passed cross-validation; five involved ages 3–5, and one included ages-3+. Two models had CVs between 0.2 and 0.5, and the rest were >0.5. Three models were consistent between the American and Canadian data sets; a comparison can be found in Table S16. American mass-at-age models were considered of low relevance to future EBFM.

Discussion

We found encouraging support for the idea that measures of 1-year-old herring in the diet of seabird chicks on colonies throughout the Gulf of Maine can provide important insight into the otherwise poorly known process of herring recruitment. Selected measures both predict recruitment to the spawning stock and illuminate spatial and temporal patterns in the first 2 years of herring life. We focus our discussion on the strongest models (Table 1), with reference, where appropriate, to models with CV between 0.2 and 0.5.

Seabirds as herring indicators

Our results show strong potential applications of seabird data to predict herring recruitment. We identified 13 seabird–fishery relationships that show great promise in predicting the herring stock. The lack of an independent recruitment index makes it impossible to predict future catch until herring reach age-2, when they can be captured by fixed gear. Bottom trawls catch few juveniles unless cohorts are strong (NEFSC 2012), and recruitment has been weak since 2005, aside from the strong 2009 age-1 cohort and potentially the 2012 cohort (Libby and Yuen 2013; Deroba 2015). Seabird diet data may thus fill an important knowledge gap in herring stock assessment and inform managers about the behavior of young herring at shorter time scales between assessments.

Seabird diets generally tracked both strong and weak age-1 cohorts as identified in stock assessments (Fig. 3); all three of the cohorts >30 000 t (1995, 2009, and 2012) and two of the four cohorts <10 000 t (2000, 2001, 2005, 2007) were reflected in chick diet at one or more colonies. These data suggest that seabirds are better at corroborating good cohorts, according to the fishery, than weak ones; our data suggest that cohorts are not equally strong across the entire region, and some of the discrepancies between the fishery and the seabirds may be related to spatial scales of the respective data sets (see below). Fishery data did a poor job of predicting weak cohorts according to seabird data; 2004, 2010, and 2011 were poor herring years for most seabirds.
yet these years are unremarkable according to the American stock assessment model. The fishery may be worse at detecting weak cohorts because technological aids allow catches even when availability may be low; five of 19 seabird–colony pairs delivered no herring to chicks in 2005, and zero abundance is unlikely to be replicated in the fishery.

Information on diet of herring predators is frequently listed as a priority for future herring research (e.g., Libby and Yuen 2013), and we present multiple measures of seabird diets located throughout the Gulf of Maine as a contribution to this requirement for EBFM. We caution, however, that herring predator data should not be used simply to corroborate findings from previous stock assessments. One of our most striking findings was that some of the strongest correlations (five of our top-ranked 13 models) were between seabird diet data and fishery data not usually incorporated into stock assessments, such as acoustic surveys and masses at age. Acoustic surveys are no longer performed on Georges Bank (M. Jech, NEFSC, Woods Hole, Massachusetts, personal communication, 2016) and were excluded from the 2012 stock assessment because they disagreed with bottom-trawl and fishery monitoring data (NEFSC 2012), yet we found a strong relationship between patterns of this acoustic survey and tern diet at MSI, which showed a marked decline after 2000 that has not otherwise been explained.

Furthermore, mass-at-age data, although considered stable by TRAC and NEFSC, actually show steep declines after 1980, as also reported by DFO (Fig. 4). Although these data may not match conclusions from landings or bottom-trawl surveys, we encourage fisheries scientists to consider the biological relevance of these data sources, especially as they support trends observed in diet of herring predators. Seabirds are often overlooked in studies of herring predators because the relative proportion of the herring stock that they consume is much less than that of groundfish or marine mammals (NEFSC 2012), but herring can be integral to seabird reproductive success and survival (Massias and Becker 1990; Durant et al. 2003; Breton and Diamond 2014), emphasizing the need for EBFM in this ecosystem. Correlations between mobile-gear landings and the data of common terns, puffins, and razorbills suggest that these seabirds select herring preferentially, further supporting the need for cautious herring management.

Spatial relationships among seabird colonies

We observed unexpected spatial patterns among seabird colonies. The easternmost colony, MSI, frequently had relationships with spawning components and fishery-derived data not shown by other seabird colonies. Western colonies often had similar lags among colonies, sometimes shared with central colonies, but never with MSI. Western common tern data sometimes showed similar patterns with data for central common terns, but more often showed patterns similar to data for central alcids, suggesting that different species may target herring differently.

There were also distinct spatial differences among western, central, and eastern colonies with regard to the abundance of herring in their seabird diets. Three of the four easternmost colonies (Matinicus Rock, Seal Island NWR, and MSI) tended to have low proportions of herring in the seabird diet, especially after 2000. Petit Manan Island, while part of this eastern group, is located close inshore and tended to have more abundant herring in the diet in the mid-2000s, but herring sharply declined in seabird diets after 2009. In contrast, chick diets in western colonies and the nearshore Eastern Egg Rock in the central Gulf of Maine showed increases in herring in recent years. Strong cohorts also tended to be spatially restricted (Fig. 3); strong cohorts in 2009, 2012, and 2014 were generally found in either the eastern or western half of the Gulf of Maine and were poor to average on the other half. These broad differences support the idea that spawning components are distinct and self-sustaining (Berkeley et al. 2004; Overholtz et al. 2004) and also suggest that juvenile herring in the eastern Gulf of Maine are now less abundant than those in the west. Links between seabirds at MSI and Georges Bank acoustic data suggest that recruitment from this spawning component may be reduced from previous years, especially after 2000. The strong relationship between the Canadian spawning component and the western seabird data corroborates descriptions in Tupper et al. (1998) that young herring occupy discrete areas, where environmental conditions may differ considerably (Fig. 6; see below). Applying a single estimate of spawning stock biomass and recruitment to all herring in the stock may put the weaker stock components at risk of overfishing (Tupper et al. 1998). The Gulf of Maine has an east–west oceanographic profile (Jordan et al. 2010; Friedland et al. 2015), with more favourable conditions for growth in the west (Tupper et al. 1998), suggesting that productivity of young herring in the east may be reduced.

Although separate quotas are allocated to different parts of the American side of the fishery (Libby and Yuen 2013), the general outlook for the most recent American stock survey was optimistic (Deroba 2015) and did not mention weak recruitment in the east. Our results suggest that more rigorous estimates for smaller regions within the Gulf of Maine would be beneficial for herring management. Although American models represent a much larger proportion of the stock relative to Canadian models, American estimates of recruitment correlated with fewer seabird colonies. A potential issue with the American model is that it combines a spatially broad data set into a single estimate of recruitment, ignoring natural oceanographic differences between Georges Bank and the eastern and western Gulf of Maine. Even within management units, the stock is divided into the northern (Gulf of Maine, SY) and southern (Georges Bank, SZ) halves of the region and does not reflect the differences in seabird diet that we have observed in the east–west pattern. Strong spatial patterns in our seabird–herring relationships imply similar patterns in distribution and numbers of juvenile herring, arguing for consideration of more spatially explicit assessments of Gulf of Maine herring.

Connectivity between spawning components and juvenile habitat

Managers have identified the connectivity between different spawning components of herring as a major source of uncertainty (Tupper et al. 1998; Overholtz et al. 2004). Our results show clear spatial patterns of juvenile herring movement and may help to resolve some of this uncertainty. Acoustic surveys by DFO track spawning components in southwestern Nova Scotia and show that larvae either remain in this vicinity or advect into the Bay of Fundy (Stephenson et al. 2015). However, we observed a consistent relationship between western and central seabird diets and the Bay of Fundy acoustic surveys, all with +3 lag. This pattern of movement is suggestive of movements of larvae from New Brunswick and western Maine (Tupper et al. 1998). Herring can first spawn at age-3, so this lag could represent the movement of juvenile herring to these breeding grounds and potentially indicate natal fidelity. A link between Canadian spawning grounds and western seabird colonies was further supported by the Canadian recruitment and mass-at-age-1 data (Fig. 6). These data sets also showed strong ties to these western and central seabird colonies, supporting the idea that a considerable portion of these Canadian spawning components produce juveniles that spend their first summers in the western and central Gulf of Maine. Curiously, although data from MSI, which is closest to the Bay of Fundy, had significant lags with the acoustic data, they were exclusively short lags of +1 year or no lag. This suggests that the juveniles around the central and western seabird colonies returned to the Bay of Fundy only just before spawning and may have experienced similar environmental conditions to those experienced by juveniles consumed by seabirds at MSI. These data may represent broad spatial movements of multiple age classes of these spawning
Fig. 6. Links between Canadian fishery data and seabird diets in the Gulf of Maine and Bay of Fundy region, based on cross-validated regression models. Two types of fishery data represented the age-1 cohort (age-1 recruitment, mass-at-age-1, upper panel), and one was linked to older cohorts (acoustic survey, lower panel). The blue and orange arrows indicate links between age-1 data and seabird colonies; the blue arrow indicates a significant relationship with both sets of data, whereas the orange arrow shows a link with just one set of data. The black arrows show ties back to the Bay of Fundy acoustic data 3 years later, suggesting movement back to Canada 3 years later. Spawning areas are shown in blue polygons, where acoustic surveys are performed. [Colour online.]
components, revealed without the need for direct assessment through tagging studies.

We observed a second unexpected spatial pattern between the Georges Bank acoustic survey and the tern diets at MSI; our post hoc analysis revealed that alcids at MSI showed a similar pattern, but only prior to 2005. These relationships were among the strongest in the model set. MSI terns often consumed >80% herring by mass in their diet in the 1990s, peaking in 1997, but herring declined in chick diets after 2000. This decline predicts the decline in biomass observed in acoustic surveys on Georges Bank in 2001, which remained at much lower levels thereafter. Although the biomass decline at Georges Bank is considered odd in the American stock assessment report (NEFSC 2012), it is one of the few data sets that matches the magnitude of decline in herring observed at this colony. Georges Bank juveniles are generally expected to remain on Georges Bank or travel “inshore” (Boyar et al. 1973; Anthony and Waring 1980; Reid et al. 1999; Overholtz et al. 2004), rather than spending time in the Bay of Fundy around MSI, and such behaviour has been suggested by others (Iles 1971; Anthony and Waring 1980 in Tupper et al. 1998). The absence of acoustic data after 2012 and the collapse of the MSI tern colony in 2006 make future links between these data sets currently impossible, but the tern colony has now been restored (Scopel and Diamond 2017), and this link could be explored further if acoustic sampling were resumed.

Models including the longest seabird time series usually had weaker relationships than shorter time series. Our post hoc analysis revealed that many longer seabird time series — three of which did not produce significant cross-validated models — had significant correlations for only part of series or had differing lags between different subsets of years. One potential explanation is that there were fundamental bottom-up changes in the ecosystem following the regime shift of 2000–2004 (Townsend et al. 2010; Smith et al. 2012) that altered the relationship between herring and their seabird predators. Another explanation is that if the connectivity between juvenile habitat and spawning components changes, or if the relative production from spawning components changes through time, the seabird–fishery relationships would also change and thus become weaker when viewed at longer temporal scales. The strong spatial pattern observed between the Bay of Fundy herring stock and western–central seabird colonies is consistent across colonies and corroborates larval movement observed in other studies (Tupper et al. 1998), leading us to believe that these relationships are real. Since the connectivity of these spawning components is poorly understood, linkages between seabirds and herring should continue to be monitored to ascertain potential changes in spatial relationships.

Diets of local seabird populations offer a means to sample and assess the size and abundance of juvenile herring from different spawning components. Herring from these colonies could also be obtained directly from the birds, generating samples to explore spatial dynamics with other techniques (DNA, otoliths, stable isotopes, etc.). A more sophisticated analysis of these spatially dispersed seabird colonies may allow better development of indices to predict the movement of young herring and the contribution of different spawning components to different parts of the Gulf of Maine.

Interpretation of lags

We observed three types of lags in our data: short lags (0 or +1 year), long lags (+2 to +5 years), and negative lags (~1 to ~2 years). We offer an interpretation of these three lags.

Long lags represent the fate of a herring cohort as it moves from potential prey to seabirds in its first or second years, into and through the fishery. These lags were found especially in models including acoustic surveys, mobile-gear landings, modeled spawning stock biomass, and masses-at-age. These data types typically represent older age classes of fish, supporting our interpretation. These lags are most useful in predicting future stock size or productivity.

Short lags are more difficult to interpret, especially when fishery data represent older age classes, such as landings and acoustic surveys; there is no obvious reason why 1-year-old herring should be prominent in seabird diet in the same year that fishery catches of older cohorts are high. While some short lags could represent cohort size (e.g., modeled age-1 recruitment), we suggest that most reflect local environmental conditions, such as temperature, which affect all age classes equally. Good or poor growth conditions experienced by juvenile herring near seabird colonies may also reflect the environmental conditions experienced by adults at the same time, depending upon coincidence in spatial distribution of young and adult herring.

Negative lags were observed in fixed-gear landings and masses-at-age models. In these cases, the fishery predicts the bird diets and potentially age-1 recruitment. Although mass-at-age could serve as a representation of condition and potential fecundity in herring adults (Berkeley et al. 2004), fixed-gear landings are harder to interpret. Fixed-gear landings capture 2-year-old fish primarily, which could indicate that the abundance of fish captured by this passive gear indicates the size or condition of the incoming cohort to the spawning population and subsequent reproductive potential. These relationships may be useful in developing an index of recruitment, although they are likely more complicated than the simple linear relationships that we tested.

Please see Supplementary Data S2 for a discussion of data limitations.

Relevance of stock data to EBFM

The most recent American stock assessment report (Deroba 2015) suggested that spawning stock biomass in the near future should be high, based on the modeled spawning stock biomass in 2014 being the highest ever recorded. However, 2014 modeled spawning stock biomass frequently appeared as an outlier in our cross-validation process, and American models have a severe retrospective pattern that tends to overestimate spawning stock biomass and underestimate fishing mortality (NEFSC 2012), by as much as 56% (TRAC 2009). While the high spawning stock biomass in 2014 likely reflects the strong 2009 cohort, more recent cohorts have been much weaker, especially in the east. This assessment report also suggested that the 2012 cohort was of equivalent magnitude to the 2009 cohort — the largest ever recorded in the region — and that the herring stock was consequently in strong shape. Seabird diet data from the eastern and central Gulf of Maine do not support the optimism expressed in this assessment; 22 of 33 diet variables were of two times or greater magnitude in 2009 versus 2012, 13 of which were five times or greater; only two showed a two times or greater magnitude of 2012 over 2009. Only diets of western common terns favoured the 2012 cohort over 2009, with eight of the 12 variables being higher in 2012 (4 > 2x).

Mass-at-age has been considered “stable” (TRAC 2009), but American and Canadian mass data both show a decline since the 1980s, the latter correlating well with seabird diet data. Declining masses-at-age imply a larger number of herring per unit of biomass, where smaller individuals have lower fecundity and survival (Paul 1980; Berkeley et al. 2004), leading to a negative feedback loop where spawning stock biomass continues to decline. Overholtz et al. (2004) observed that mean masses-at-age in the American stock components had declined significantly after 1980, representing a 20% decrease in biomass for fish at age-6 — of similar timing and magnitude to Canadian data — yet most recent assessments by TRAC and NEFSC have not identified this long-term decline in mass-at-age as an area of concern. The variable lag patterns observed between seabird data and Canadian masses-at-age suggest that these relationships are not resolved well by this analysis, but inclusion of positive and negative lags suggest that
both adults and juveniles are affected by this decline in mass-age.

The fixed-gear fishery — which is restricted to the eastern Gulf of Maine — has had poor landings for the past decade and showed negative lags with seabirds at three of the easternmost colonies. These trends reflect declining herring in the east and represent economic and social needs of part of the fishing community that has been overlooked. Our result is a reversal of the finding in Amey (1998), who examined the predictive value of MSI Arctic terns to the weir fishery in New Brunswick. Amey’s study was carried out during 1990–1997 when herring dominated MSI Arctic tern diet and weir catches nearby were high; different relationships can be expected when herring availability falls below the threshold when predators switch to alternative prey, as seems to have occurred since 2000 (see above). Herring data from mobile-gear landings and modeled spawning stock biomass typically correlated with central Gulf of Maine alcid and western common terns; these relationships were much weaker for central terns and eastern seabirds, suggesting that they are also getting overlooked in stock assessment. Eastern and central terns generally had stronger relationships with acoustic surveys or fixed-gear landings data, suggesting that these measures better represent herring availability to terns and may be underutilized in stock assessment. Arctic terns in central colonies were particularly underrepresented, leading to the question of which data would represent these birds’ needs.

Importance of ecological variables

Sea-surface temperatures in the Gulf of Maine have been increasing more rapidly than in most of the world’s oceans since 2004, exceeding historic values since 2012 (Pershing et al. 2015); further, regime shifts in the late 1980s and late 1990s have changed zooplankton communities substantially (Hare and Kane 2012). This argues that ecosystem-based management in the Gulf of Maine must take ecosystem changes into account, yet environmental data are not currently included in the stock assessment model (NEFSC 2012). Using temperature as an example, Bell et al. (2015) warned that changes in temperature can have effects as great as those of fishing, and Pershing et al. (2015) found a strong effect of temperature on the survival and growth of Atlantic cod (Gadus morhua) in the Gulf of Maine, leading to an underestimate of cod biomass, perpetual overfishing, and a delay in stock recovery. Herring are susceptible to changes in temperature at all life stages (Tupper et al. 1998; Rijnsdorp et al. 2009; Corten 2013); increasing temperatures are especially dangerous for young fish, raising metabolic rate without increased food to support their greater energetic needs, resulting in greater mortality and thus depressed recruitment (Brander 2007; Rijnsdorp et al. 2009). Recruitment has been shown to be sensitive to temperature in several herring stocks in the Northeast Atlantic (Cardinale et al. 2009; Ottersen et al. 2013; Payne et al. 2013). Increasing temperatures can also lead to trophic mismatch, as well as changes in behaviour, mortality rate, timing of maturity, growth rate, and generation time (Stevenson and Scott 2005; Hsieh et al. 2006; Rijnsdorp et al. 2009).

No ecological approach to fishery management can legitimately exclude environmental variables, especially when, like temperature, they are known to strongly influence fish biology; we strongly recommend the inclusion of at least temperature data in future assessment models to account for rapid environmental changes in this ecosystem.

Overall, we found strong relationships between seabird diet and a variety of sources of herring data; the most striking patterns were clear spatial differences in those relationships among colonies, which reflect spatial structuring in herring stocks and likely patterns of movements of herring in the years before they recruit into the stock and the fishery. Some types of herring data that are not routinely used in stock assessments — notably acoustic surveys and mass-at-age data — provided relationships as convincing as those with the more commonly used data types (mobile-gear landings, modeled spawning stock biomass). This suggests that these sources provide important insight into the process of recruitment, which is widely acknowledged to be poorly understood. We urge more focused incorporation of such data and data on seabird diets, as well as environmental variables such as temperature, if herring management is to move closer to a truly ecological approach. The dependence of seabirds on prerecruit life stages of herring, and the widespread and long-term collection of seabird diet data at all major seabird colonies, offers a unique opportunity for major progress toward the goal of a broader-based approach to herring management.

Acknowledgements

The authors acknowledge support for this study from the Pew Charitable Trusts and supporters of the National Audubon Society’s Seabird Restoration Program. We also acknowledge support from the Canadian Wildlife Service, Environment Canada, the New Brunswick Wildlife Trust Fund, the Natural Sciences and Engineering Research Council of Canada, and the University of New Brunswick. We thank M. Power, R. Singh, and R. Stephenson for providing helpful advice and mass-at-age data that were integrated into this document. Likewise, we thank M. Jech for providing acoustic data from Georges Bank. We acknowledge L. Welch and S. Williams for providing seabird diet data from Petit Manan Island. Statistical advice and coding help were provided by G. Dauphin, S. Edwards, A. Einfeldt, L. Jesson, and R. Malenfant. We also thank two anonymous reviewers for their helpful comments.

References


Published by NRC Research Press


U.S.A.


decht, the Netherlands.


Published by NRC Research Press