1 **Seabird diets as bioindicators of Atlantic herring recruitment and stock size: a new tool for ecosystem-**

2 **based fisheries management**

3 Lauren C. Scopel[1], Antony W. Diamond[1], Stephen W. Kress[2], Adrian R. Hards[3], and Paula Shannon [4]

4 [1] Atlantic Laboratory for Avian Research, University of New Brunswick, PO Box 4400, Fredericton, New

5 Brunswick, E3B 5A3. L.scopel@unb.ca, diamond@unb.ca

6 [2] National Audubon Society Seabird Restoration Program, 159 Sapsucker Woods Road, Ithaca, NY,

7 14850, USA swk3@cornell.edu

8 [3] University of New Brunswick, PO Box 4400, Fredericton, New Brunswick, E3B 5A3.

9 adrian.hards@unb.ca

10 [4] National Audubon Society Seabird Restoration Program, 12 Audubon Road, Bremen, ME, 04551, USA

11 pshannon@audubon.org

12

13

14 Corresponding author:

15 Lauren Scopel

16 PO Box 4400 28 Dineen Drive, University of New Brunswick, Fredericton, NB E3B 5A3

17 Cell: (506) 449-8044

18 l.scopel@unb.ca

19

20   Abstract

21

22   Ecosystem-based fishery management requires understanding of relationships between exploited fish

23   and their predators, such as seabirds.   We used exploratory regression analyses to model relationships

24   between Atlantic herring (*Clupea harengus*) in the diet of seabird chicks at nine nesting colonies in the

25   Gulf of Maine, and four types of fishery- and survey-derived herring data. We found several strong

26   relationships, which suggests spatial structuring in herring stocks and likely patterns of herring

27   movements before they recruit into the fishery. Some types of herring data seldom used in stock

28   assessments – notably acoustic surveys, fixed-gear landings, and weight-at-age – correlated as strongly

29   with seabird data as more commonly used series, such as mobile-gear landings and modeled spawning

30   stock biomass. Seabird chick diets collected at specific locations thus offer a promising means to assess

31   the size, distribution, and abundance of juvenile herring across a broad area prior to recruitment, which

32   is a major source of uncertainty in fisheries. Common terns showed the most potential as a

33   bioindicator, correlating well and showing consistent spatial patterns with 11 of 13 fishery data series.

34

35   **Introduction**

36   Fisheries managers world-wide are increasingly expected to oversee harvested stocks with due regard to

37   their interactions with other species in the ecosystem (Stephenson 2012). Successful application of this

38   approach, often referred to as ecosystem-based fishery management or EBFM, is widely recognised as

39   essential to the maintenance of healthy marine ecosystems and the communities they support

40   (Christensen et al. 1996; Dickey-Collas et al. 2013; Skern-Mauritzen et al. 2016).

41   A fundamental feature of EBFM is understanding the multiple interactions among constituent species

42   (Travis et al. 2014), including (but not limited to) harvested species and their prey. Fishery harvest is

43   treated as a key ecosystem service to be sustained, yet many other predators rely on the same prey

44   species, and sustaining their populations is an equally valid ecosystem service (Diamond 2012; Fogarty

45   2014). The idea that the needs of natural predators 'deserve' consideration alongside those of fisheries

46   is increasingly becoming accepted (Travis et al. 2014) and incorporated into both global biodiversity

47   goals (Convention on Biodiversity no date) and national fishery-management policies (Stephenson 2012;

48   Skern-Mauritzen et al. 2016).

49   Seabirds are among the most widespread and best-studied predators of marine fish, and they are

50   affected by large-scale changes in abundance of prey stocks especially as a result of over-fishing (Duffy

51   1983; Cury et al. 2011; Cook et al. 2014). Seabirds may also reflect other changes in the marine

52   environment including contamination and oceanographic and climatic change, leading to well-

53   established use of 'seabirds as environmental indicators' or SEI (Cairns 1987; Diamond and Devlin 2003;

54   Einoder et al. 2009). Most of the SEI literature examines various behavioral responses – such as time

55   budgets (Jodice et al. 2006) or reproductive success (Monaghan et al. 1989; Wanless et al. 2005) – to

56   changing food availability, but Parsons et al. (2008) pointed out that seabird chick diet data offer "the

57   most immediate signal of availability," which are not widely available because they are time-consuming

58   to collect. We take advantage of our long time-series of seabird chick diet data to conduct exploratory

59   analyses using data routinely collected to manage one widely-harvested forage fish, Atlantic herring

60   *Clupea harengus,* which has been prominent in the diet of seabirds and many other top predators (see

61   below).

62   Forage fish are the target of many fisheries throughout the world (Pikitch et al. 2012). They are generally

63   small or medium-sized, pelagic, schooling species, often with high fat content (Harris and Hislop 1978;

64   Hislop et al. 1991; Wanless et al. 2005), characterised ecologically by their position mid-way in the food

65   chain between phytoplankton and larger predators; in such marine systems the energy required by

66   higher trophic levels is channeled through the forage fish (Pikitch et al. 2012). In such 'wasp-waist'

67   systems (Bakun et al. 2009) the forage fish play a 'keystone' role (i.e. a species that exerts influences out

68   of proportion to its abundance, *sensu* Paine 1995; Johnson et al. 2017), supporting a variety of top

69   predators including larger fish, marine mammals, and marine birds; if these forage fish are the target of

70   a fishery, it is critically important to avoid over-fishing because of all the predators that rely on the same

71   prey as the fishery (Cury et al. 2011). If such forage fish show high caloric density (i.e., are high in fat

72   content) they are particularly likely to be targeted by optimal foragers and function as keystone prey.

73   In systems where seabirds feed on forage fish which are also the target of a fishery, there is enhanced

74   potential both for negative impacts of the fishery on the birds and for the birds to respond quickly to

75   changes in the fish stock; much of the SEI literature refers to such systems. Among many examples are:

76   sandeels (or sandlance) *Ammodytes* spp. in the North Sea (Monaghan et al. 1989; Frederiksen et al.

77   2008; Cook et al. 2014) and Gulf of Alaska (Jodice et al. 2006; Suryan et al. 2006); capelin *Mallotus*

78   *villosus* around Newfoundland and Labrador (Montevecchi and Myers 1996; Regehr and Montevecchi

79   1997), the Barents Sea (Barrett and Furness 1990) and Gulf of Alaska (Jodice et al. 2006); anchovy and

80   sardines in the Benguela Current (Crawford et al. 2008; Green et al. 2015), the Humboldt Current (Duffy

81   1983) and California Current (Mills et al. 2007; Sydeman et al. 2009; Thompson et al. 2012); and Atlantic

82   herring *Clupea harengus* in Iceland, Norway (Anker-Nilssen 1987), and the Gulf of Maine (Diamond and

83   Devlin 2003; Breton and Diamond 2014).

84   Atlantic herring in the Gulf of Maine play a widely recognised keystone role (Kenney et al. 1997; Bakun

85   et al. 2009; Diamond 2012) as a forage fish. Predators with a predominantly herring diet include bluefin

86   tuna (*Thunnus thynnus*), themselves the object of an important fishery (Golet et al. 2007; 2015),

87   humpback *Megaptera novaeangliae* and fin whales *Balaenoptera physalus,* other cetaceans and seals

88   (Kenney 1997), and several species of long-distance migratory seabirds as well as those breeding locally

89   (Diamond 2012). However, different predators consume different sizes (and therefore age classes) of

90   the herring stock. Juvenile herring hatched in fall are referred to as "0-group" until 1 January when they

91   become 1-group (or 1-yr-olds). Herring reach sexual maturity at 3-4 years but can enter the fishery at

92   age 2 (Stephenson et al. 1993; Fig. 1). Herring in the Gulf of Maine spawn mostly in late summer or fall,

93   so those eaten by seabirds breeding the following summer are most often 1 yr old, though the larger

94   auks (razorbill *Alca torda* and common murre *Uria aalge*) also take 2-yr-old fish. Seabirds sample herring

95   younger than 2 yrs old, offering the possibility that data on their diets could predict recruitment, and aid

96   in stock assessment and EBFM.

97   Our focus in this paper is on assessing possible relationships between the herring content in the diet of

98   seabird chicks and various measures of herring abundance, with a view to identifying possible predictors

99   of future herring stock size. We assume that seabirds take preferred prey (in this case, 1-yr-old herring)

100   in relation to the abundance of that prey; thus, a high proportion of herring in the diet reflects an

101   abundance of herring within the foraging range of the colony. Specifically, we assess relationships

102   between sets of herring data commonly used in stock assessments in both Canadian and American parts

103   of the Gulf of Maine, and several measures of diets of breeding seabirds and their chicks. From this

104   assessment, we examine the potential of diet data from seabird species and colonies as indicators of

105   value to future EBFM. We include diet data from multiple species at widely-separated colonies, allowing

106   us to address both the different spatial and temporal scales at which seabirds and fisheries operate, and

107   the varied foraging behaviours of different species of seabird (Greenstreet et al. 2000; Becker and

108   Chapdelaine 2003). Our analyses are driven by two broad hypotheses: 1) the diets of seabirds indicate

109   the strength of a herring cohort at age 1, which predicts future stock size (seabird diets predict fishery),

110   and 2) data on breeding herring from fisheries predict future productivity, and the prevalence of herring

111   in future seabird diets (fishery predicts seabird diets). To reflect these hypotheses, and the temporal

112   disjuncture between the age-1 cohorts targeted by seabirds and the older ages targeted by various parts

113   of the fishery, we perform an exploratory analysis examining relationships between fishery and seabird

114   data at lags ranging from -3 (fishery predicts seabirds 3 years later) to +5 (seabirds predict fishery 5

115   years later). We identify the strongest seabird-fishery relationships, rank herring data for their

116   relevance to herring predators, and identify seabird colonies and species as potentially useful indicators

117   of the herring stock.

118

119   **Methods**

120   *Seabird study area and diet data collection*

121   Seabird diet data were collected from nine colonies in the Gulf of Maine and Bay of Fundy region, where

122   long-term seabird monitoring programs have been conducted since the early 1970s (Fig. 2). Seven

123   Maine colonies (Stratton Island, Outer Green Island, Jenny Island, Pond Island National Wildlife Refuge

124   (NWR), Eastern Egg Rock, Matinicus Rock, and Seal Island NWR) are managed by National Audubon

125   Society's Seabird Restoration Program (SRP), one is managed by the U.S. Fish and Wildlife Service

126   (USFWS; Petit Manan Island, Maine), and one (Machias Seal Island, New Brunswick, "MSI") is managed

127   by the Canadian Wildlife Service and monitored by the University of New Brunswick (UNB). Protocols to

128   collect diet data were similar across islands. The animal care committee at UNB reviewed annual

6

129   procedures performed at MSI, and all SRP and USFWS data collection was in accordance with federal

130   banding permits. We consider four colonies as "western" (Stratton Island, Outer Green Island, Jenny

131   Island, and Pond Island NWR), three colonies as "central" (Eastern Egg Rock, Matinicus Rock, and Seal

132   Island NWR), and two as "eastern" (Petit Manan Island, MSI).

133   We included diet data from four species – two surface-feeding tern species (common tern *Sterna*

134   *hirundo* and Arctic tern *S. paradisaea*) and two diving alcid species (Atlantic puffin *Fratercula arctica* and

135   razorbill *Alca torda*). Common terns breed on all nine colonies, and we used data from Arctic terns on

136   four (Matinicus Rock, Seal Island, Petit Manan Island, and MSI), Atlantic puffins on four (Eastern Egg

137   Rock, Matinicus Rock, Seal Island, and MSI), and razorbills on two (Matinicus Rock and MSI). The latter

138   three species tended to nest in colonies in the central and eastern parts of the Gulf of Maine,

139   representing the southern limit of their North American ranges.

140   Feeding watches were performed throughout the chick-rearing period of each study species on each

141   island, from multiple observation blinds erected in each colony. Observers used binoculars to identify

142   common prey species at each colony, estimating prey length relative to the adult's bill length. For each

143   feeding, observers recorded the nest number, prey recipient, number of prey, prey length, and prey

144   species. Further details on seabird diet observations are included in Supplementary Data S1.

145   A full summary of the time series for all species at all colonies is included in Table S1. Time series at SRP

146   colonies begin in 1991 or 1992 for Arctic terns, 2002 or 2003 for common terns, and 2005 for alcids.

147   Time series at Petit Manan Island begin in 2000 for Arctic terns and 2002 for common terns. At MSI,

148   data begin in 1990 for Arctic terns and 1995 for common terns and alcids. The tern colony at MSI

149   collapsed in 2006 (Diamond 2009; Gaston et al. 2009), so no tern feeding observations could be

150   performed from 2006-2013. The diet study for Arctic terns resumed in 2014 with the restoration of the

7

151   colony (Scopel and Diamond 2017), and diet estimates for common terns were derived from video

152   observations in 2016.

153

154   *Seabird diet data treatment*

155   Seabird diet data treatment follows that performed in Scopel and Diamond (*in press*). All raw feeding

156   data were converted into annual estimates of wet weight. For each species at each colony, taxa that

157   made up at least 1% of the diet by number were retained for conversion into weight (Table S2). If

158   possible, length-weight conversion formulae were developed from prey samples collected

159   opportunistically by UNB at MSI as part of their annual protocol (available at

160   http://www.unb.ca/research/alar/msi-seabirds/protocol-methods.html); prey samples were measured

161   for total length to the nearest 0.1 mm, and wet weight to the nearest 0.1 g. For prey taxa with too few

162   samples to develop conversions (n = 8), we used values from the literature (Table S3). Once all prey

163   conversions were developed, we applied the conversion formulae to our annual diet data for each

164   colony-species pair (e.g., common tern-Jenny Island, Atlantic puffin-Seal Island), to derive estimates of

165   wet weight for each prey taxon and for each year. Sums of weight for each prey taxon were tallied and

166   summarized for each year.

167   For this study, we calculated four representations of herring in the seabird diet: two rates (herring mass

168   per feeding, herring mass per observed chick-hour (i.e., the total number of hours each chick was

169   observed at a nest (tern) or plot-hour (alcid))), proportion of herring in the diet, and average mass of

170   individual herring. A summary of these representations of seabird diet, including sample sizes, can be

171   found in Tables S4-S12 in supplemental material. Strength of herring cohort by colony-species pair is

172   included in Table 1.

173

174   *Herring stock data sources*

175   Herring in the Gulf of Maine and Bay of Fundy region are managed by an American and a Canadian

176   fishery, respectively; although both countries manage their quotas and stocks independently, they do

177   share data, and issue status reports jointly through the Transboundary Resource Assessment Committee

178   (TRAC). The Canadian fishery manages the Bay of Fundy and Scotian Shelf herring as the 4WX herring

179   stock complex, requiring the separation of Bay of Fundy herring (4X) from reported totals in the 4VWX

180   reports. The American fishery identifies the Gulf of Maine and Georges Bank as separate spawning

181   components (5Y and 5Z, respectively), but both are modeled as part of the same stock complex (NEFSC

182   2012).

183   Owing to the diversity of data that can represent a fish stock, we considered a variety of potential

184   herring data types for this analysis. We selected four broad categories that represent different aspects

185   of the herring fishery: fishery-independent acoustic surveys of spawning stock biomass, landings,

186   weights-at-age, and output from stock assessment models. Although landings are relatively simple to

187   measure and data are abundant, they do represent targeted and technologically-assisted sampling of

188   fish that does not necessarily represent the behaviour of small piscivores, such as seabirds. We

189   therefore included data from acoustic surveys to obtain a fishery-independent estimate of spawning

190   stock biomass. The Canadian fishery has reported declining average weights-at-age since the 1970s

191   (Department of Fisheries and Oceans (DFO) 2015), so we included Canadian weight data to reflect this

192   potentially important property of the stock. Finally, we included two types of output from stock

193   assessment models. Age-1 recruitment was estimated by both the American and Canadian stock

194   assessments via forward-cast (Age Structured Assessment Program with a Beverton-Holt stock-

195   recruitment relationship) and back-cast (Virtual Population Assessment, no longer used officially) age-

196   structured models, respectively; these serve as the best available data to estimate the abundance of 1-

197   yr-old herring in the fishery, which can be compared directly with seabird diets. Finally, we included

198   estimates of spawning stock biomass from the American stock assessment model, for comparison with

199   landings and acoustic surveys.

200   Landings data were obtained from stock assessment reports. We included mobile-gear landings

201   (primarily purse-seine and mid-water trawl) from the American fishery (Deroba 2015), mobile-gear

202   landings from New Brunswick (Stephenson et al. 1998; Singh et al. 2016), combined landings (American

203   mobile-gear landings and all fixed-gear landings, Deroba 2015), and fixed-gear landings from

204   southwestern New Brunswick (Deroba 2015). Most fixed-gear landings in the region are caught in weirs

205   in southwestern New Brunswick.

206   Acoustic data were obtained for the Canadian and American stocks; Canadian data were obtained from

207   stock assessment reports (DFO 2015), and American data were obtained directly from M. Jech at the

208   NEFSC, following the format published in the American stock assessment report (NEFSC 2012). Canadian

209   surveys sampled herring at Trinity Ledge, German Bank, and Scots Bay, Nova Scotia; American surveys

210   sampled herring at Georges Bank (Fig. 2). Acoustic survey data in both regions begin in 1999, and

211   continue through 2012 in the USA and 2014 in Canada.

212   Weight-at-age data were obtained directly from M. Power at DFO. We selected three types of weight

213   data – weight-at-age estimates for age-1 fish (the youngest available, the same age that most of the

214   seabirds would consume), fish aged 3-5 years (which represent the majority of the spawning stock), and

215   fish aged 3+ years. The majority of spawners in the stock are 3-5 years old, but the entire spawning

216   component can reach age 10+. We compared the importance of just young spawners (3-5 yrs) and the

217   entire spawning stock; although younger spawners are more abundant, older and larger fish may

218   contribute more to future recruitment owing to greater fecundity (Wootton 1998; Hsieh et al. 2006).

219   DFO also identified a significant decline in the average growth of herring between ages 2 and 3 in recent

220   years (M. Power, DFO, St. Andrews, New Brunswick, personal communication, 2016), which we included

221   as a fourth data type in our weight category as "growth increment".

222   Finally, we included spawning stock biomass estimates from the American model (Deroba 2015), and

223   estimates of the number of age-1 herring as predicted by the American (NEFSC 2012) and Canadian

224   (obtained directly from R. Singh, DFO) models.

225   Complete herring data used in this analysis, including years of available data, are included in table S13-

226   S14 in supplemental material.   Strength of herring cohorts, as measured by the American stock

227   assessment model, are included in Table 1.

228

229   *Data analysis – Exploratory assessment of lags*

230   Herring sampled by seabirds and by the fishery are rarely of the same age class; most seabirds consume

231   1- or 2-year old herring, but the herring fishery targets 2+ (fixed gear) and 3+ year old (mobile gear,

232   acoustic surveys) fish (Fig. 1). This age discrepancy offers the potential to predict future stock data from

233   the younger age classes consumed by seabirds. In order to investigate potentially useful predictive lags,

234   we identified significant lags using the cross-correlation function (ccf(), package "stats") in R (version

235   3.1.2, R Core Team 2014). We examined all lags between -3 to +5 for each seabird-colony and fishery

236   data pair; positive lags indicate that seabird diets predict fishery data, and vice versa for negative lags.

237   Any lag that was deemed significant by the cross-correlation function (>0.4, R Core Team 2014) was

238   retained for further analysis. We identified three classes of lags: (1) no lag or lags of +1 years, "no" or

239   "short lags", (2) lags of +2 to +5 years, "long lags", and (3) lags of -1 to -3 years, "negative lags."

240   We used 19 colony-species pairs of seabird diet data (e.g., common tern on Matinicus Rock, Atlantic

241   puffin at Eastern Egg Rock), each with four representations of seabird-derived herring (e.g., herring

242   weight per feeding), and correlated them to 13 variables derived from the herring fishery (e.g., fixed-

243   gear landings). For average mass of individual herring (seabird), we correlated these to the four fishery-

244   derived weight variables only.

245

246   *Data analysis – Validation of significant correlations*

247   Following the exploratory assessment of lags, we examined all significant seabird-fishery correlations in

248   greater detail. All analyses were performed in R. All data were normalized (mean of 0, standard

249   deviation +/-1). We used univariate lagged regression to predict fishery-derived estimates of herring

250   from seabird-derived estimates of herring. Owing to the direct comparison of two herring measures, we

251   tested only simple linear relationships; we assumed that differences in cohort size would persist linearly

252   throughout the time series.

253   Since seabirds can consume 1- or 2-year-old herring, we tested the possibility of temporal

254   autocorrelation by including an autoregressive (AR) correlation structure within our models (Zuur et al.

255   2009). We prepared two sets of models for each significant lag, one with and one without an AR1

256   correlation structure. We used generalized least squares regression (GLS, gls(), package "nlme",

257   Pinheiro et al. 2016) and compared each model pair using AICc (package "AICcmodavg", Mazerolle

258   2015); for each model, the structure with the lowest AICc was selected.

259   Models were built with fishery data as the response variable and seabird data as the predictor. Series of

260   models were produced for each fishery variable and ranked using AICc. Any models that ranked lower

261   than a model using only year as a predictor (i.e., the "null" model) were considered uninformative and

262   removed from further consideration.

263   After fitting GLS models, we checked models for meaningful effect sizes and explained variation, using

264   regression slopes and McFadden's pseudo-R^2 (McFadden 1974), respectively. Models that scored lower

265   than 0.5 in either category were considered uninformative and removed from further consideration. In

266   some cases, a single seabird-fishery lag was represented multiple times with different rate and

267   proportional data; if one model was ranked more than 3 ΔAICc units higher than another model within

268   the same seabird-fishery lag (e.g., Seal Island puffins-acoustic +3), we removed it from consideration. All

269   instances with lags within 3 ΔAICc units were kept.

270   For all remaining models, leave-one-out cross-validation (Efron and Tibshirani 1993) was performed to

271   assess the model's predictive ability and stability of slope. Models where slopes dropped below 0.25

272   were considered uninformative and removed from further consideration. Once cross-validation was

273   performed, we reduced the model set to a single instance of each seabird-fishery pair, if possible, based

274   on prediction error (CV), pseudo-R^2, slope, and potential outlier years. Multiple lags for one seabird-

275   fishery pair were retained if they were well supported.   Models with greater pseudo-R^2, greater slopes,

276   and lower prediction error were considered stronger.

277

278   *Relevance rankings*

279   GLS models were separated into potential quantitatively and qualitatively useful lags; CVs of <0.5 were

280   considered suitable for quantitative prediction, but those >0.5 were considered useful to represent

281   general qualitative trends only. CVs <0.2 were considered the strongest predictive models, and these

282   were considered potentially reliable bioindicators.

283   Using the cross-validated GLS models, fishery data were ranked based on their potential relevance to

284   future EBFM. Ranking criteria for each fishery data type included the number of seabird species, seabird

285   colonies, and geographical subregions (western, central, or eastern) that were represented by cross-

286   validated models, the number of models that passed cross-validation, and the number of models with

287   CVs <0.2. Model sets that were representative of a larger spatial area, highly correlated with more

288   species, and had strong predictive value were considered the most relevant to future study.

289

290   *Post hoc analyses*

291   Following our initial analyses, we performed a *post hoc* comparison between our seabird variables and

292   American weight-at-age data (NEFSC 2012). We found a marked reduction in weight-at-age after age 3

293   in both Canadian and U.S. data (Fig. 3) which encouraged us to include both datasets, using the same

294   age categories (age 1, 3-5, and 3+) as the Canadian weight data. The American time series included data

295   through 2011.

296   We also examined the importance of time series length within our longest seabird time series. We

297   divided four time series (MSI Atlantic puffins and razorbills, Matinicus Rock and Seal Island NWR Arctic

298   terns) into two subsets – an "early" time series including years 2005 and earlier (to emulate the length

299   of the MSI tern series), and a "late" time series including years 2002 and later (to emulate the length of

300   the SRP common tern series). We examined cross-correlations between these subsets and compared

301   them with cross-correlations of the full time series, in order to determine if changes in the seabird-

302   fishery relationship over time weakened the strength of the observed correlations in these data. This

303   division coincides with the timing of a regime shift between 2000 and 2004 (Smith et al. 2012) with

304   potential consequences for nutrient status and the plankton community (Townsend et al. 2010) which

305   can be expected to affect responses in species higher in the food-web such as forage fish and their

306   predators.

307

308   **Results**

309   *Assessment of lags and validation of models*

310   We performed 6,555 seabird-fishery cross-correlation comparisons, identifying 472 seabird-fishery data

311   pairs with significant lags during the first stage of the analysis; 71 included an autoregressive correlation

312   structure. We performed cross-validation on 347 GLS models, retaining 70 (Table S15). Of these, 19 had

313   CVs >0.5, 38 had CVs between 0.2 and 0.5, and 13 had CVs <0.2. Models with CVs <0.2 we consider to

314   be the highest quality (Table 2), and are examined in full detail below. Models with CVs between 0.2

315   and 0.5 had quantitative value but showed higher prediction error, and those with CVs >0.5 were

316   considered qualitatively useful only; general trends from these two lower ranked categories are also

317   discussed below.

318

319   *Top ranked models – CV <0.2*

320   A complete list of the 13 top-ranked models, including their CVs, pseudo-R^2, and slopes, can be found in

321   Table 2. Depictions of these models can also be found in Figs. 4 and S1-S2 in supplemental material.

322   Eight of 13 fishery data types and 10 of 19 seabird-colony pairs were involved in this model set,

323   including all four seabird species, and six of nine seabird colonies, encompassing all three geographic

324   subregions. Acoustic surveys, landings, and model output were equally represented in this model set,

325   but only one model included weight data. Common tern models were the most abundant, featured nine

326   times, and six colonies of common terns were represented in this model set. Eight models had long lags,

327   four models had short or no lag, and just one model had a negative lag.

328

329   *Trends in lower-ranked models – 0.2 < CV < 0.5 and CV > 0.5*

330   Twelve of 13 fishery data types were represented in these models; only mobile-gear landings from New

331   Brunswick were absent from cross-validated models. Several fishery data types featured prominently in

332   these lower ranked models, including estimated recruitment from the Canadian fishery, Canadian

333   weight-at-age data for all three age categories, modeled spawning stock biomass, and mobile-gear

334   landings from the American fishery.

335   All nine seabird colonies were included in these models, although only 16 of 19 seabird-colony pairs

336   were represented. Razorbills from MSI, and Arctic terns from Matinicus Rock and Seal Island NWR were

337   not part of any of the cross-validated models; no strong relationships could be found between fishery

338   data and any of these seabirds. Of the remaining seabird-colony pairs, common terns were part of 25

339   models, puffins in 17, Arctic terns in 9, and razorbills in 6.

340   Negative lags featured in 17 models, particularly those including weight, fixed- or mobile-gear landings

341   data. Long lags were in 12 models, including acoustic surveys, mobile-gear landings, spawning stock

342   biomass, and weight data. The remaining 28 models included short or no lag, and incorporated 11

343   different fishery variables.

344

345   *By fishery data type – all models*

346   Acoustic surveys from the Bay of Fundy showed long lags with western common tern and central alcid

347   data, and short or no lags with terns at MSI and razorbills at Matinicus Rock. Acoustic surveys from

348   Georges Bank linked exclusively to terns at MSI (Fig. 4), at long and short or no lags.

349   Fixed-gear landings had a consistent lag of -2 for four of six cross-validated models; three seabird

350   species from MSI, both tern species from Petit Manan Island, and common terns from Matinicus Rock

351   showed links to fixed-gear landings, which primarily occur in waters near MSI (Fig. S1). American

352   mobile-gear landings correlated with data from seven seabird-colony pairs, spread across western and

353   central colonies, but neither eastern nor Arctic tern data were included in these models. Lags for

354   mobile-gear models included three long lags, two short lags, and two negative lags. American mobile-

355   gear landings had the second-highest number of significant correlations in early model-testing, but

356   many models failed cross-validation owing to high-leverage points in 2006 and 2009.

357   Modeled spawning stock biomass was the most prolific fishery type in the full cross-validated model set,

358   including 10 seabird-colony pairs at eight colonies, yet many models failed cross-validation owing to

359   high leverage from the 2014 data point. Five models included long lags, and the remaining models had

360   short or no lags. Recruitment data from the USA were most strongly linked to common tern and puffin

361   data from two central colonies, but overall were much more spatially limited; all models had no lag.

362   This indicates that recruits, as defined by the American stock assessment, are of the same age as the fish

363   consumed by seabirds. Conversely, recruits from the Canadian stock assessment were represented

364   primarily by +1-yr lags, indicating that recruits, as defined by the Canadian stock assessment, are one

365   year younger than the fish consumed by seabirds. Canadian recruitment models involved common

366   terns, puffins, and razorbills from five central or western colonies. Although Canadian recruitment data

367   were more widespread in their links with seabird data relative to American recruitment data, American

368   models on average had lower prediction error; no Canadian recruitment models were included in the

369   top-ranked model set. Common terns and puffins at Seal Island NWR were the only seabird-colony pairs

370   to correlate with both sets of recruitment data.

371   Canadian weight data were included in 23 total models, but only one ranked highly and lags were

372   inconsistent, especially in the weight-at-age 1 category. Similarly to the Canadian recruitment data,

373   weight-at-age 1 models involved common tern, puffin, and razorbill data from central and western

374   colonies, incorporating five colonies. The lags featured in the weight-at-age 1 data, however, were quite

375   variable, including seven negative lags and four long lags. Weight-at-age 3-5 data and growth increment

376   were included in seven models, representing alcids from central colonies and terns at Petit Manan

377   Island. Weight-at-age 3+ data were linked with terns at MSI only.

378

379   *By seabird species and colony – all models*

380   Common terns were represented strongly; all nine colonies had models that passed cross-validation,

381   and eight colonies were included in models with CVs <0.5. Common tern models also involved 11 of 13

382   fishery data series, showing high concordance with a variety of fishery data types. The strongest models

383   for common terns included both acoustic surveys, all three types of landings, and American recruitment

384   and spawning stock biomass estimates. Common tern data also formed the strongest models, as ranked

385   by CV and by pseudo-R^2.

386   Arctic tern data at MSI were linked with all four types of fishery data, including five variables, most often

387   by +1 lags. MSI Arctic tern data correlated most frequently and most strongly with acoustic survey data,

388   especially in Georges Bank. At Petit Manan Island, Arctic tern data were linked to just three variables,

389   representing either weight or landings data, and relationships were weaker than the MSI models. Arctic

390   tern data from central colonies did not correlate strongly with anything.

391   Atlantic puffins at all four colonies had models with CVs <0.5, including nine fishery variables, yet only

392   one model was included in the top ranked set, featuring weight-at-age data. Puffins at Seal Island NWR

393   and Matinicus Rock were the best represented in these models, showing strong relationships with

394   weight, landings, and model data. All central colonies had models including Canadian recruitment

395   estimates, American mobile-gear landings, and Canadian weight-at-age data. Puffins at MSI were linked

396   only to fixed-gear landings.

397   Razorbills at Matinicus Rock were linked to six fishery variables, including all four types of fishery data.

398   Only one model was included in the top-ranked set, incorporating estimates of spawning stock biomass.

399   Razorbill models were similar to puffin models from central colonies, also including Canadian

400   recruitment estimates, mobile-gear landings, and weight-at-age data. Razorbills at MSI showed no

401   relationship with any fishery data.

402

403   *Relevance rankings*

404   Rankings are listed in Table 3. Of the 13 fishery data series, two have high relevance to future

405   ecosystem-based fisheries management, four have moderate relevance, six have low or limited

406   relevance, and one has poor relevance.

407

408   *Post hoc analyses*

409   Six American weight-at-age models passed cross-validation; five involved ages 3-5, and one included

410   ages 3+. Two models had CVs between 0.2 and 0.5, and the rest were >0.5. Three models were

411   consistent between the American and Canadian datasets; a comparison can be found in Table S16.

412   American weight-at-age models were considered of low relevance to future EBFM.

413   We compared the four truncated seabird time series with 13 fishery data types, for a total of 52

414   comparisons. In 25 cases, one subset had a significant lag while the other subset had none; in 12 cases,

415   both subsets had significant lags, but the lag lengths were two or more years apart between subsets.

416   Acoustic biomass models had significant correlations prior to 2005 only, and model-derived recruitment

417   had stronger correlations in 2002 and later. Weights and landings had cases where early and late lags

418   were the only ones with significant correlations, depending on the variable.

419

420   **Discussion**

421   We found encouraging support for the idea that measures of 1-yr-old herring in the diet of seabird

422   chicks on colonies throughout the Gulf of Maine can provide important insight into the otherwise

423   poorly-known process of herring recruitment. Selected measures both predict recruitment to the

424   spawning stock, and illuminate spatial and temporal patterns in the first two years of herring life. We

425   focus our discussion on the strongest models (Table 2), with reference, where appropriate, to models

426   with CV between 0.2 and 0.5.

427

428   *Seabirds as herring indicators*

429   Our results show strong potential applications of seabird data to predict herring recruitment. We

430   identified 13 seabird-fishery relationships that show great promise in predicting the herring stock. The

431   lack of an independent recruitment index makes it impossible to predict future catch until herring reach

432   age 2, when they can be captured by fixed gear. Bottom trawls catch few juveniles unless cohorts are

433   strong (NEFSC 2012), and recruitment has been weak since 2005, aside from the strong 2009 age-1

434   cohort, and potentially the 2012 cohort (Libby and Yuen 2013; Deroba 2015). Seabird diet data may

435   thus fill an important knowledge gap in herring stock assessment, and inform managers about the

436   behavior of young herring at shorter time scales between assessments.

437   Seabird diets generally tracked both strong and weak age-1 cohorts as identified in stock assessments

438   (Table 1); all three of the cohorts >30,000 tons (1995, 2009, and 2012), and two of the four <10,000 tons

439   (2000, 2001, 2005, 2007), were reflected in chick diet at one or more colonies. These data suggest that

440   seabirds are better at corroborating good cohorts, according to the fishery, rather than weak ones; our

441   data suggest that cohorts are not equally strong across the entire region, and some of the discrepancies

442   between the fishery and the seabirds may be related to spatial scales of the respective datasets (see

443   below). Fishery data did a poor job of predicting weak cohorts according to seabird data; 2004, 2010,

444   and 2011 were poor herring years for most seabirds, yet these years are unremarkable according to the

445   American stock assessment model. The fishery may be worse at detecting weak cohorts because

446   technological aids allow catches even when availability may be low; five of 19 seabird-colony pairs

447   delivered no herring to chicks in 2005, and zero abundance is unlikely to be replicated in the fishery.

448   Information on diet of herring predators is frequently listed as a priority for future herring research (e.g.,

449   Libby and Yuen 2013), and we present multiple measures of seabird diets located throughout the Gulf of

450   Maine as a contribution to this requirement for EBFM. We caution, however, that herring predator data

451   should not be used to simply corroborate findings from previous stock assessments. One of our most

452   striking findings was that some of the strongest correlations (five of our top-ranked 13 models) were

453   between seabird diet data and fishery data not usually incorporated into stock assessments, such as

454   acoustic surveys and weights-at-age. Acoustic surveys are no longer performed on Georges Bank (M.

455   Jech, NEFSC, Woods Hole, MA, personal communication, 2016), and were excluded from the 2012 stock

456   assessment because they disagreed with bottom-trawl and fishery monitoring data (NEFSC 2012), yet

457   we found a strong relationship between patterns of this acoustic survey and tern diet at MSI, which

458   showed a marked decline after 2000 that has not otherwise been explained.

459   Furthermore, weight-at-age data, although considered stable by TRAC and NEFSC, actually show steep

460   declines after 1980, as also reported by DFO (Fig. 3). Although these data may not match conclusions

461   from landings or bottom-trawl surveys, we encourage fisheries scientists to consider the biological

462   relevance of these data sources, especially as they support trends observed in diet of herring predators.

463   Seabirds are often overlooked in studies of herring predators because the relative proportion of the

464   herring stock that they consume is much less than that of groundfish or marine mammals (NEFSC 2012),

465   but herring can be integral to seabird reproductive success and survival (Massias and Becker 1990;

466   Durant et al. 2003; Breton and Diamond 2014), emphasizing the need for EBFM in this ecosystem.

467   Correlations between mobile-gear landings and data of common terns, puffins, and razorbills suggest

468   that these seabirds select herring preferentially, further supporting the need for cautious herring

469   management.

470

471   *Spatial relationships among seabird colonies*

472   We observed unexpected spatial patterns among seabird colonies. The easternmost colony, MSI,

473   frequently had relationships with spawning components and fishery-derived data not shown by other

474   seabird colonies. Western colonies often had similar lags among colonies, sometimes shared with

475   central colonies, but never with MSI. Western common tern data sometimes showed similar patterns

476   with central common terns, but more often showed patterns similar to central alcids, suggesting that

477   different species may target herring differently.

478   There were also distinct spatial differences among western, central, and eastern colonies with regard to

479   the abundance of herring in their seabird diets. Three of the four easternmost colonies (Matinicus Rock,

480   Seal Island NWR, and MSI) tended to have low proportions of herring in the seabird diet, especially after

481   2000. Petit Manan Island, while part of this eastern group, is located close inshore and tended to have

482   more abundant herring in the diet in the mid-2000s, but herring sharply declined in seabird diets after

483   2009. In contrast, chick diets in western colonies and the nearshore Eastern Egg Rock in the central Gulf

484   of Maine showed increases in herring in recent years. Strong cohorts also tended to be spatially

485   restricted (Table 1); strong cohorts in 2009, 2012, and 2014 were generally found in either the eastern

486   or western half of the Gulf of Maine, and were poor to average on the other half. These broad

487   differences support the idea that spawning components are distinct and self-sustaining (Berkeley et al.

22

488   2004; Overholtz et al. 2004), and also suggest that juvenile herring in the eastern Gulf of Maine are now

489   less abundant than those in the west. Links between seabirds at MSI and Georges Bank acoustic data

490   suggest that recruitment from this spawning component may be reduced from previous years, especially

491   after 2000. The strong relationship between the Canadian spawning component and the western

492   seabird data corroborates descriptions in Tupper et al. (1998) that young herring occupy discrete areas,

493   where environmental conditions may differ considerably (Fig. 5; see below). Applying a single estimate

494   of spawning stock biomass and recruitment to all herring in the stock may put the weaker stock

495   components at risk of overfishing (Tupper et al. 1998). The Gulf of Maine has an east-west

496   oceanographic profile (Jordaan et al. 2010; Friedland et al. 2015), with more favourable conditions for

497   growth in the west (Tupper et al. 1998), suggesting that productivity of young herring in the east may be

498   reduced.

499   Although separate quotas are allocated to different parts of the American side of the fishery (Libby and

500   Yuen 2013), the general outlook for the most recent American stock survey was optimistic (Deroba

501   2015), and did not mention weak recruitment in the east. Our results suggest that more rigorous

502   estimates for smaller regions within the Gulf of Maine would be beneficial for herring management.

503   Although American models represent a much larger proportion of the stock relative to Canadian models,

504   American estimates of recruitment correlated with fewer seabird colonies. A potential issue with the

505   American model is that it combines a spatially broad dataset into a single estimate of recruitment,

506   ignoring natural oceanographic differences between Georges Bank and the eastern and western Gulf of

507   Maine. Even within management units, the stock is divided into the northern (Gulf of Maine, 5Y) and

508   southern (Georges Bank, 5Z) halves of the region, and does not reflect the differences in seabird diet

509   that we have observed in an east-west pattern. Strong spatial patterns in our seabird-herring

510   relationships imply similar patterns in distribution and numbers of juvenile herring, arguing for

511   consideration of more spatially explicit assessments of Gulf of Maine herring.

512

513   *Connectivity between spawning components and juvenile habitat*

514   Managers have identified the connectivity between different spawning components of herring as a

515   major source of uncertainty (Tupper et al. 1998; Overholtz et al. 2004). Our results show clear spatial

516   patterns of juvenile herring movement, and may help to resolve some of this uncertainty. Acoustic

517   surveys by DFO track spawning components in southwestern Nova Scotia, and show that larvae either

518   remain in this vicinity, or advect into the Bay of Fundy (Stephenson et al. 2015). However, we observed

519   a consistent relationship between western and central seabird diets and the Bay of Fundy acoustic

520   surveys, all with +3 lag. This pattern of movement is suggestive of movements of larvae from New

521   Brunswick and eastern Maine (Tupper et al. 1998). Herring can first spawn at age 3, so this lag could

522   represent the movement of juvenile herring to these breeding grounds, and potentially indicate natal

523   fidelity. A link between Canadian spawning grounds and western seabird colonies was further

524   supported by the Canadian recruitment and weight-at-age 1 data (Fig. 5). These datasets also showed

525   strong ties to these western and central seabird colonies, supporting the idea that a considerable

526   portion of these Canadian spawning components produce juveniles that spend their first summers in the

527   western and central Gulf of Maine. Curiously, although data from MSI, which is closest to the Bay of

528   Fundy, had significant lags with the acoustic data, they were exclusively short lags of +1 year or no lag.

529   This suggests that the juveniles around the central and western seabird colonies returned to the Bay of

530   Fundy only just before spawning, and may have experienced similar environmental conditions to those

531   experienced by juveniles consumed by seabirds at MSI. These data may represent broad spatial

532   movements of multiple age classes of these spawning components, revealed without the need for direct

533   assessment through tagging studies.

534   We observed a second unexpected spatial pattern between the Georges Bank acoustic survey and the

535   tern diets at MSI; our post hoc analysis revealed that alcids at MSI showed a similar pattern, but only

536   prior to 2005. These relationships were among the strongest in the model set. MSI terns often

537   consumed >80% herring by mass in their diet in the 1990s, peaking in 1997, but herring declined in chick

538   diets after 2000. This decline predicts the decline in biomass observed in acoustic surveys on Georges

539   Bank in 2001, which remained at much lower levels thereafter. Although the biomass decline at

540   Georges Bank is considered odd in the American stock assessment report (NEFSC 2012), it is one of the

541   few datasets that matches the magnitude of decline in herring observed at this colony. Georges Bank

542   juveniles are generally expected to remain on Georges Bank or travel "inshore" (Boyar et al. 1973;

543   Anthony and Waring 1980; Reid et al. 1999; Overholtz et al. 2004), rather than spending time in the Bay

544   of Fundy around MSI, and such behaviour has been suggested by others (Iles 1971; Anthony and Waring

545   1980 *in* Tupper et al. 1998). The absence of acoustic data after 2012 and the collapse of the MSI tern

546   colony in 2006 make future links between these datasets currently impossible, but the tern colony has

547   now been restored (Scopel and Diamond 2017) and this link could be explored further if acoustic

548   sampling were resumed.

549   Models including the longest seabird time series usually had weaker relationships than shorter time

550   series. Our post hoc analysis revealed that many longer seabird time series – three of which did not

551   produce significant cross-validated models – had significant correlations for only part of series, or had

552   differing lags between different subsets of years. One potential explanation is that there were

553   fundamental bottom-up changes in the ecosystem following the regime shift of 2000-2004 (Townsend

554   et al. 2010; Smith et al. 2012) that altered the relationship between herring and their seabird predators.

555   Another explanation is that if the connectivity between juvenile habitat and spawning components

556   changes, or if the relative production from spawning components changes through time, the seabird-

557   fishery relationships would also change, and thus become weaker when viewed at longer temporal

558   scales. The strong spatial pattern observed between the Bay of Fundy herring stock and western-central

559   seabird colonies is consistent across colonies and corroborates larval movement observed in other

560   studies (Tupper et al. 1998), leading us to believe that these relationships are real. Since the

561   connectivity of these spawning components is poorly understood, linkages between seabirds and

562   herring should continue to be monitored to ascertain potential changes in spatial relationships.

563   Diets of local seabird populations offer a means to sample and assess the size and abundance of juvenile

564   herring from different spawning components. Herring from these colonies could also be obtained

565   directly from the birds, generating samples to explore spatial dynamics with other techniques (DNA,

566   otoliths, stable isotopes, etc.). A more sophisticated analysis of these spatially dispersed seabird

567   colonies may allow better development of indices to predict the movement of young herring and the

568   contribution of different spawning components to different parts of the Gulf of Maine.

569

570   *Interpretation of lags*

571   We observed three types of lags in our data – short (0 or +1 years), long (+2 to +5 years), and negative

572   lags (-1 to -2 years). We offer an interpretation of these three lags.

573   Long lags represent the fate of a herring cohort as it moves from potential prey to seabirds in its first or

574   second years, into and through the fishery. These lags were found especially in models including

575   acoustic surveys, mobile-gear landings, modeled spawning stock biomass, and weights-at-age. These

576   data types typically represent older age classes of fish, supporting our interpretation. These lags are

577   most useful in predicting future stock size or productivity.

578   Short lags are more difficult to interpret, especially when fishery data represent older age classes, such

579   as landings and acoustic surveys; there is no obvious reason why 1-yr-old herring should be prominent in

580   seabird diet in the same year that fishery catches of older cohorts are high. While some short lags could

581   represent cohort size (e.g., modeled age-1 recruitment), we suggest that most reflect local

582   environmental conditions, such as temperature, which affect all age classes equally. Good or poor

583   growth conditions experienced by juvenile herring near seabird colonies may also reflect the

584   environmental conditions experienced by adults at the same time, depending upon coincidence in

585   spatial distribution of young and adult herring.

586   Negative lags were observed in fixed-gear landings and weight-at-age models. In these cases, the

587   fishery predicts the bird diets, and potentially age-1 recruitment. Although weight-at-age could serve as

588   a representation of condition and potential fecundity in herring adults (Berkeley et al. 2004), fixed-gear

589   landings are harder to interpret.   Fixed-gear landings capture 2-yr-old fish primarily, which could

590   indicate that the abundance of fish captured by this passive gear indicates the size or condition of the

591   incoming cohort to the spawning population, and subsequent reproductive potential. These

592   relationships may be useful in developing an index of recruitment, although they are likely more

593   complicated than the simple linear relationships that we tested.

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

595

596   *Relevance of stock data to EBFM*

597   The most recent American stock assessment report (Deroba 2015) suggested that spawning stock

598   biomass in the near future should be high, based on the modeled spawning stock biomass in 2014 being

599   the highest ever recorded. However, 2014 modeled spawning stock biomass frequently appeared as an

600   outlier in our cross-validation process, and American models have a severe retrospective pattern that

601   tends to overestimate spawning stock biomass and underestimate fishing mortality (NEFSC 2012), by as

602   much as 56% (TRAC 2009). While the high spawning stock biomass in 2014 likely reflects the strong

603   2009 cohort, more recent cohorts have been much weaker, especially in the east. This assessment

604   report also suggested that the 2012 cohort was of equivalent magnitude to the 2009 cohort – the largest

605   ever recorded in the region – and that the herring stock was consequently in strong shape. Seabird diet

606   data from the eastern and central Gulf of Maine do not support the optimism expressed in this

607   assessment; 22 of 33 diet variables were of 2x or greater magnitude in 2009 vs. 2012, 13 of which were

608   5x or greater; only two showed a 2x or greater magnitude of 2012 over 2009. Only diets of western

609   common terns favoured the 2012 cohort over 2009, with 8 of the 12 variables being higher in 2012 (4 >

610   2x).

611   Weight-at-age has been considered "stable" (TRAC 2009), but American and Canadian weight data both

612   show a decline since the 1980s, the latter correlating well with seabird diet data. Declining weights-at-

613   age imply a larger number of herring per unit of biomass, where smaller individuals have lower

614   fecundity and survival (Pauly 1980; Berkeley et al. 2004), leading to a negative feedback loop where

615   spawning stock biomass continues to decline. Overholtz et al. (2004) observed that mean weights-at-

616   age in the American stock components had declined significantly after 1980, representing a 20%

617   decrease in biomass for fish at age 6 – of similar timing and magnitude to Canadian data – yet most

618   recent assessments by TRAC and NEFSC have not identified this long-term decline in weight-at-age as an

619   area of concern. The variable lag patterns observed between seabird data and Canadian weights-at-age

620   suggest that these relationships are not resolved well by this analysis, but inclusion of positive and

621   negative lags suggest that both adults and juveniles are affected by this decline in weight-at-age.

622   The fixed-gear fishery – which is restricted to the eastern Gulf of Maine – has had poor landings for the

623   past decade, and showed negative lags with seabirds at three of the easternmost colonies. These trends

624   reflect declining herring in the east, and represent economic and social needs of part of the fishing

625   community that has been overlooked. Our result is a reversal of the finding in Amey (1998), who

626   examined the predictive value of MSI Arctic terns to the weir fishery in New Brunswick. Amey's study

627   was carried out 1990-97 when herring dominated MSI Arctic tern diet and weir catches nearby were

628   high; different relationships can be expected when herring availability falls below the threshold when

629   predators switch to alternative prey, as seems to have occurred since 2000 (see above). Herring data

630   from mobile-gear landings and modeled spawning stock biomass typically correlated with central Gulf of

631   Maine alcids and western common terns; these relationships were much weaker for central terns and

632   eastern seabirds, suggesting that they are also getting overlooked in stock assessment. Eastern and

633   central terns generally had stronger relationships with acoustic surveys or fixed-gear landings data,

634   suggesting that these measures better represent herring availability to terns, and may be underutilized

635   in stock assessment. Arctic terns in central colonies were particularly underrepresented, leading to the

636   question of which data would represent these birds' needs.

637

638   *Importance of ecological variables*

639   Sea-surface temperatures in the Gulf of Maine have been increasing more rapidly than in most of the

640   world's oceans since 2004, exceeding historic values since 2012 (Pershing et al. 2015); further, regime

641   shifts in the late 1980s and late 1990s have changed zooplankton communities significantly (Hare and

642   Kane 2012). This argues that ecosystem-based management in the Gulf of Maine must take ecosystem

643   changes into account, yet environmental data are not currently included in the stock assessment model

644   (NEFSC 2012). Using temperature as an example, Bell et al. (2015) warned that changes in temperature

645   can have effects as great as those of fishing, and Pershing et al. (2015) found a strong effect of

646   temperature on the survival and growth of Atlantic cod (*Gadus morhua*) in the Gulf of Maine, leading to

647   an overestimate of cod biomass, perpetual overfishing, and a delay in stock recovery. Herring are

648   susceptible to changes in temperature at all life stages (Tupper et al. 1998; Rijnsdorp et al. 2009; Corten

649   2013); increasing temperatures are especially dangerous for young fish, raising metabolic rate without

650   increased food to support their greater energetic needs, resulting in greater mortality and thus

651   depressed recruitment (Brander 2007; Rijnsdorp et al. 2009). Recruitment has been shown to be

652   sensitive to temperature in several herring stocks in the northeast Atlantic (Cardinale et al. 2009;

653   Ottersen et al. 2013; Payne et al. 2013). Increasing temperatures can also lead to trophic mismatch, as

654   well as changes in behaviour, mortality rate, timing of maturity, growth rate, and generation time

655   (Stevenson and Scott 2005; Hsieh et al. 2006; Rijnsdorp et al. 2009). No ecological approach to fishery

656   management can legitimately exclude environmental variables, especially when, like temperature, they

657   are known to strongly influence fish biology; we strongly recommend the inclusion of at least

658   temperature data in future assessment models in order to account for rapid environmental changes in

659   this ecosystem.

660

661   Overall we found strong relationships between seabird diet and a variety of sources of herring data; the

662   most striking patterns were clear spatial differences in those relationships among colonies, which reflect

663   spatial structuring in herring stocks and likely patterns of movements of herring in the years before they

664   recruit into the stock and the fishery. Some types of herring data which are not routinely used in stock

665   assessments– notably acoustic surveys and weight-at-age data – provided relationships as convincing as

666   those with the more commonly used data types (mobile-gear landings, modeled spawning stock

667   biomass). This suggests that these sources provide important insight into the process of recruitment,

668   which is widely acknowledged to be poorly understood. We urge more focused incorporation of such

669   data, and data on seabird diets, as well as environmental variables such as temperature, if herring

670   management is to move closer to a truly ecological approach. The dependence of seabirds on pre-recruit

671   life stages of herring, and the widespread and long-term collection of seabird diet data at all major

672   seabird colonies, offers a unique opportunity for significant progress toward the goal of a broader-based

673   approach to herring management.

685   **References**

686   Amey, K. 1998. Seabirds as indicators of changes in availability and commercial weir landings of herring.

687   M.Sc. thesis, Department of Biology, University of New Brunswick, Fredericton, NB.

688   Anker-Nilssen, T. 1987. The breeding performance of Puffins *Fratercula arctica* on Røst, Northern

689   Norway, 1979-1985. Fauna norvegica Series C, Cinclus **10**: 21-38.

690   Anthony, V.C., and Waring, G.T. 1980. Estimate of herring spawning stock biomass and egg-production

691   for the Georges Bank-Gulf of Maine Region. Northeast Fisheries Science Center, Northwest Atl. Fish.

692   Organ. Sci. Counc. Res. Doc. 1980b;80/1X/135 Ser. No. N209.

693   Bakun, A., Babcock, E.A., and Santora, C. 2009. Regulating a complex adaptive system via its wasp-waist:

694   grappling with ecosystem-based management of the New England herring fishery. ICES J. Mar. Sci. **66**(8):

695   1768–1775. doi:10.1093/icesjms/fsp073.

696   Barrett, R.T., and Furness, R.W. 1990. The prey and diving depths of seabirds on Hornøy, North Norway

697   after a decrease in the Barents Sea capelin stocks.   Ornis Scand. **21**(3): 179-186. Available from

698   https://www.jstor.org/stable/3676777 [accessed 10 January 2017].

699   Becker, P., and Chapdelaine, G. 2003. Further development of seabird monitoring. *In* Seabirds as

700   monitors of the marine environment. *Edited by* Tasker, M.L. and Furness, R.W. ICES Cooperative

701   Research Report No. 258. 52-60.

702   Bell, R.J., Richardson, D.E., Hare, J.A., Lynch, P.D., and Fratantoni, P.S. 2015. Disentangling the effects of

703   climate, abundance, and size on the distribution of marine fish: an example based on four stocks from

704   the Northeast US shelf. ICES J. Mar. Sci. **72**(5): 1311–1322. doi:10.1093/icesjms/fsu217.

705   Berkeley, S.A., Hixon, M.A., Larson, R.J., and Love, M.S. 2004. Fisheries sustainability via protection of

706   age structure and spatial distribution of fish populations. Fisheries **29**(8): 23-32. doi:10.1577/1548-

707   8446(2004)29[23:FSVPOA]2.0.CO;2.

708   Boyar, H.C., Marak, R.R., Perkins, F.E., and Clifford, R.A. 1973. Seasonal distribution and growth of larval

709   herring (Clupea harengus L.) in the Georges Bank-Gulf of Maine area from 1962 to 1970. J. Cons., Cons.

710   Perm. Int. Explor. Mer **35**(1): 36-51. doi:10.1093/icesjms/35.1.36.

711   Brander, K.M. 2007. Global fish production and climate change. Proc. Natl. Acad. Sci. U. S. A. **104**(50):

712   19709-19714. doi:10.1073/pnas.070205.

713   Breton, A.R., and Diamond, A.W. 2014. Annual survival of adult Atlantic Puffins *Fratercula arctica* is

714   positively correlated with Herring *Clupea harengus* availability. Ibis **156(**1): 35–47.

715   doi:10.1111/ibi.12100.

716   Cairns, D.K. 1987. Seabirds as indicators of marine food supplies. Biological Oceanography **5**(4): 261–

717   271. doi:10.1080/01965581.1987.10749517.

718 Cardinale, M., Möllmann, C., Bartolino, V., Casini, M., Kornilovs, G., Raid, T., Margonski, P., Grzyb, A.,

719 Raitanemi, J., Gröhsler, T., and Flinkman, J. 2009. Effect of environmental variability and spawner

720 characteristics on the recruitment of Baltic herring *Clupea harengus* populations. Mar. Ecol. Prog. Ser.

721 **388**:221-234. https://doi.org/10.3354/meps08125

722 Christensen, N.L., Bartuska, A.M., Brown, J.H., Carpenter, S., D'Antonio, C., Francis, F., Franklin, J.F.,

723 MacMahon, J.A., Noss, F.R., Parsons, D.J., Peterson, C.H., Turner, M.G., and Woodmansee, R.G. 1996.

724 Report of the Ecological Society Committee on the scientific basis for ecosystem management. Ecol.

725 Appl. **6**(3): 665-691. doi:10.2307/2269460.

726 Convention on Biological Diversity. Nd. Generic and specific indicators for assessing progress in the

727 attainment of the Aichi biodiversity targets, including an assessment of their main characteristics

728 [online]. Available from *https://www.cbd.int/doc/strategic-plan/strategic-plan-indicators-en.pdf*

729 [accessed 5 December 2016].

730 Cook, A.S.C.P., Dadam, D., Mitchell, I., Ross-Smith, V.H., and Robinson, R.A. 2014. Indicators of seabird

731 reproductive performance demonstrate the impact of commercial fisheries on seabird populations in

732 the North Sea. Ecol. Indic. **38**: 1–11. doi:10.1016/j.ecolind.2013.10.027.

733 Corten, A. 2013. Recruitment depressions in North Sea herring. ICES J. Mar. Sci. **70**(1): 1-15.

734 doi:10.1093/icesjms/fss187.

735 Crawford, R.J.M., Sabarros, P.S., Fairweather, T., Underhill, L.G., and Wolfaard, A.C. 2008. Implications

736 for seabirds off South Africa of a long-term change in the distribution of sardine. Afr. J. Marine Sci **30**(1):

737 177–184*.* doi:10.2989/AJMS.2008.30.1.18.468.

738 Cury, P.M., Boyd, I.L., Bonhommeau, S., Anker-Nilssen, T., Crawford, R.J.M., Furness, R.W., Mills, J.A.,

739 Murphy, E.J., Österblom, H., Paleczny, M., Piatt, J.F., Roux, J.-P., Shannon, L., and Sydeman, W.J. 2011.

740 Global seabird response to forage fish depletion—one-third for the birds. Science **334**(6063): 1703–

741 1706. doi:10.1126/science.1212928.

742 Department of Fisheries and Oceans. 2015. 2015 Assessment of 4VWX herring. DFO Can. Sci. Advis.

743 Sec. Res. Doc. 2015/040.

744 Department of Fisheries and Oceans. 2016. 4VWX herring 2016 update report. DFO Can. Sci. Advis.

745 Sec. Sci. Resp. 2016/036.

746 Deroba, J. 2015. Atlantic Herring Operational Assessment Report 2015. U.S. Dept. Commer., Northeast

747 Fish. Sci. Cent. Ref. Doc. 15-16.

748 Diamond, A.W. 2009. Vicissitudes of the tern colony on Machias Seal Island. NB Naturalist **36**(3): 116-

749 117.

750 Diamond, A.W. 2012. Managing for migrants. *In* Advancing an ecosystem approach in the Gulf of

751 Maine. *Edited by*: Stephenson, R.L., Annala, J.H., Runge, J.A., and Hall-Arber, M. Am. Fish. Soc. Symp. 79.

752 Amer. Fisheries Soc., Bethesda, MD. Pp. 311-320.

753 Diamond, A.W., and Devlin, C.M. 2003. Seabirds as indicators of changes in marine ecosystems:

754 Ecological monitoring on Machias Seal Island. Environ. Monit. Assess. **88**(1): 153–175.

755 doi:10.1023/A:1025560805788.

756 Dickey-Collas, M., Engelhard, G.H., Rindorf, A., Raab, K., Smout, S., Aarts, G., van Deurs, M., Brunel, T.,

757 Hoff, A., Lauerburg R.A.M., Garthe, S., Haste Andersen, K., Scott, F., van Kooten, T., Beare, D., and Peck,

758 M.A. 2013. Ecosystem-based management objectives for the North Sea: riding the forage fish

759 rollercoaster. ICES J. Mar. Sci. **71**(1): 128-142. doi:10.1093/icesjms/fst075.

760 Duffy, D.C. 1983. Environmental uncertainty and commercial fishing: effects on Peruvian guano birds.

761 Biological Conservation **26**(3): 227-238. doi:10.1016/0006-3207(83)90075-7.

762   Durant, J.M., Anker-Nilssen, T., and N.C. Stenseth. 2003. Trophic interactions under climate fluctuations:

763   the Atlantic Puffin as an example. Proceedings: Biological Sciences **270** (1523): 1461-1466.

764   Efron, B., and R. Tibshirani. 1993. An introduction to the bootstrap. Monographs on Statistics and

765   Applied Probability 57. Chapman and Hall, Boca Raton, FL.

766   Einoder, L.D. 2009. A review of the use of seabirds as indicators in fisheries and ecosystem

767   management. Fisheries Research **95**(1): 6–13. doi:10.1016/j.fishres.2008.09.024.

768   Fogarty, M.J. 2014. The art of ecosystem-based fishery management. Can. J. Fish. Aquat. Sci. **71**:479-

769   490.

770   Frederiksen, M., Jensen, H., Daunt, F., Mavor, R. A., and Wanless, S. 2008. Differential effects of a local

771   industrial sand lance fishery on seabird breeding performance. Ecol. Appl. **18**(3): 701–710.

772   doi:10.1890/07-0797.1.

773   Friedland, K.D., Leaf, R.T., Kane, J., Tommasi, D., Asch, R.G., Rebuck, N., Ji, R., Large, S.I., Stock, C., and

774   Saba, V.S. 2015. Spring bloom dynamics and zooplankton biomass response on the US Northeast

775   Continental Shelf. Cont. Shelf Res. **102**: 47-61. doi:10.1016/j.csr.2015.04.005.

776   Gaston, A.J., Bertram, D.F., Boyne, A.W., Chardine, J.W., Davoren, G., Diamond, A.W., Hedd, A.,

777   Montevecchi, W.A., Hipfner, J.M., Lemon, M.J.F., Mallory, M.L., Rail, J.-F., and Robertson, G.J. 2009.

778   Changes in Canadian seabird populations and ecology since 1970 in relation to changes in oceanography

779   and food webs. Environ. Rev. **17**: 267–286. doi:10.1139/A09-013.

780   Golet, W.J., Cooper, A.B., Campbell, R. and Lutcavage, M. 2007. Decline in condition of northern bluefin

781   tuna (*Thunnus thynnus*) in the Gulf of Maine. Fish. Bull. **105**: 390-395.

782 Golet, W.J., Record, N.R., Lehuta, S., Lutcavage, M., Glauradi, B., Campbell, R., and Pershing, A.J. 2015.

783 The paradox of the pelagics: why bluefin tuna can go hungry in a sea of plenty. Mar. Ecol. Progr. Ser.

784 **527**: 181-192.

785 Green, D.B., Klages, N.T.W., Crawford, R.J.M., Coetzee, J.C., Dyer, B.M., Rishworth, G.M., and Pistorius,

786 P.A. 2015. Dietary change in Cape gannets reflects distributional and demographic shifts in two South

787 African commercial fish stocks. ICES J. Mar. Sci. **72**(3): 771–781. doi:10.1093/icesjms/fsu203.

788 Greenstreet, S.P.R., Becker, P.H., Barrett, R.T., Fossum, P., and Leopold, M F. 2000. Consumption of pre-

789 recruit fish by seabirds and the possible use of this as an indicator of fish stock recruitment. *In* Diets of

790 seabirds and consequences of changes in food supply. *Edited by* Furness, R.W., and Tasker, M.L. ICES

791 Cooperative Research Report no. 232. 6-17.

792 Hare, J.A., and Kane, J. 2012. Zooplankton in the Gulf of Maine – a changing perspective. Am. Fish. Soc.

793 Symp. **79**:115-135.

794 Harris, M.P., and Hislop, J.R.G. 1978. The food of young puffins *Fratercula arctica*. J. Zool. Lond. **185**:213-

795 236.

796 Hislop, J.R.G., Harris, M.P., and Smith, J.G.M. 1991. Variation in the calorific value and total energy

797 content of the lesser sandeel (*Ammodytes marinus*)   and other fish preyed on by seabirds. J. Zool. Lond.

798 **224**:501-517.

799 Hsieh, C., Reiss, C.S., Hunter, J.R., Beddington, J.R., May, R.M., and Sugihara, G. 2006. Fishing elevates

800 variability in the abundance of exploited species. Nature **443**: 859-862. doi:10.1038/nature05232.

801 Iles, T.D. 1971. The retention inside the Bay of Fundy of herring larvae spawned off the southwest coast

802 of Nova Scotia. Selected papers from the 1971 annual meeting. International Commission for the

803 Northwest Atlantic Fisheries Redbook Part **3**: 93-103.

804   Jodice, P.G.R., Roby, D.D., Turco, K.R., Suryan, R.M., Irons, D.B., Piatt, J.F., Shultz, M.T., Roseneau, D.G.,

805   Kettle, A.B., and Anthony, J.A. 2006. Assessing the nutritional stress hypothesis: Relative influence of

806   diet quantity and quality on seabird productivity. Mar. Ecol.: Prog. Ser. **325**: 267–279.

807   doi:10.3354/meps325267.

808   Johnson, S.A., Ober, H.K., and Adams, D.C. 2017. Are keystone species effective umbrellas for habitat

809   conservation? A spatially explicit approach. Journal for Nature Conservation 37:47–55.

810   Jordaan, A., Chen, Y., Townsend, D.W., and Sherman, S. 2010. Identification of ecological structure and

811   species relationships along an oceanographic gradient in the Gulf of Maine using multivariate analysis

812   with bootstrapping. Can. J. Fish. Aquat. Sci. **67**: 701-719.

813   Kenney, R., Scott, G., Thompson, T., and Winn, H. 1997. Estimates of prey consumption and trophic

814   impacts of cetaceans in the USA northeast continental shelf ecosystem. J. Northw.

815   Atl. Fish. Sci. **22**: 155–171. doi: 10.1.1.555.4576.

816   Libby, D., and Yuen, M. 2013. Review of the fishery management plan for Atlantic Sea Herring (*Clupea*

817   *harengus harengus*) for the 2012 fishing year. Atlantic States Marine Fisheries Commission.

818   Massias, A., and Becker, P.H. 1990. Nutritive value of food and growth in Common Tern *Sterna hirundo*

819   chicks. Ornis Scand. **21**(3): 187–194. Available from https://www.jstor.org/stable/3676778 [accessed 10

820   January 2017].

821   Mazerolle, M.J. 2015. AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c)

822   [online]. R package version 2.0-3. http://CRAN.R-project.org/package=AICcmodavg [accessed 1 June

823   2014].

824   McFadden, D. 1974. Conditional logit analysis of qualitative choice behaviour. *In* Frontiers in

825   econometrics. *Edited by* Zarembka, P. Academic Press, New York. NY. Pp. 105-142.

826   Mills, K.L., Laidig, T., Ralston, S., and Sydeman, W.J. 2007. Diets of top predators indicate pelagic juvenile

827   rockfish (*Sebastes* spp.) abundance in the California Current System. Fisheries Oceanography **16**(3): 273–

828   283. doi:10.1111/j.1365-2419.2007.00429.x.

829   Monaghan, P., Uttley, J.D., and Okill, J.D. 1989. Terns and sandeels: seabirds as indicators of

830   changes in marine fish populations. J. Fish Biol. Suppl. A*,* **35**: 339-40. doi:10.1111/j.1095-

831   8649.1989.tb03080.x.

832   Montevecchi, W.A., and Myers, R.A. 1996. Dietary changes of seabirds reflect shifts in pelagic food

833   webs. Sarsia **80**: 313-322. doi:10.1080/00364827.1996.10413606.

834   Northeast Fisheries Science Center. 2012. 54th Northeast regional stock assessment workshop (54th

835   SAW) assessment report. U.S. Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 12-18.

836   Ottersen, G., Stige, L.C., Durant, J.M., Chan, K.S., Rouyer, T.A., Drinkwater, K.F., and Stenseth N.C. (2013)

837   Temporal shifts in recruitment dynamics of North Atlantic fish stocks: effects of spawning stock and

838   temperature. Mar. Ecol. Prog. Ser. **480**:205-225. https://doi.org/10.3354/meps10249

839   Overholtz, W.J., Jacobson, L.D., Melvin, G.D., Cieri, M., Power, M., Libby, D., and Clark K. 2004. Stock

840   Assessment of the Gulf of Maine-Georges Bank Atlantic Herring complex, 2003. U.S. Dept. Commer.,

841   Northeast Fish. Sci. Cent. Ref. Doc. 04-06.

842   Paine, R.T. 1995. A conversation on refining the concept of keystone species. Conservation Biology 9:

843   962-964.

844   Parsons, M., Mitchell, I., Butler, A., Ratcliffe, N., Frederiksen, M., Foster, S., and Reid, J.B. 2008. Seabirds

845   as indicators of the marine environment. ICES J. Mar. Sci. **65**(8): 1520–1526.

846   doi:10.1093/icesjms/fsn155.

847   Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters, and mean

848   environmental temperature in 175 fish stocks. ICES J. Mar. Sci. **39**(2): 175-192.

849   doi:10.1093/icesjms/39.2.175.

850   Payne, M.R., Stine, D.R., Worsøe, Clausen L., Munk, P., Mosegaard, H., and Nash, R.D.M. (2013)

851   Recruitment decline in North Sea herring is accompanied by reduced larval growth rates. Mar. Ecol.

852   Prog. Ser. **489**:197-211. https://doi.org/10.3354/meps10392

853   Pershing, A.J., Alexander, M.A., Hernandez, C.M., Kerr, L.A., Le Bris, A., Mills, K.E., Nye, J.A., Record, N.R.,

854   Scannell, H.A., Scott, J.D., Sherwood, G.D., and Thomas, A.C. 2015. Slow adaptation in the face of rapid

855   warming leads to collapse of the Gulf of Maine cod fishery. Science **350**(6262): 809-

856   812. doi:10.1126/science.aac9819.

857   Pikitch, E., Boersma, P.D., Boyd, I.L., Conover, D.O., Cury, P., Essington, T., Heppell, S.S., Houde, E.D.,

858   Mangel, M., Pauly, D., Plaganyi, E., Sainsbury, K. and Steneck, R.S. 2012. Little fish, big impact: managing

859   a crucial link in ocean food webs. Lenfest Ocean Program, Washington, DC.

860   Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team. 2016. nlme: Linear and nonlinear mixed

861   effects models [online]. R package version 3.1-127. Available from http://CRAN.R-

862   project.org/package=nlme [accessed 15 November 2016].

863   R Core Team. 2014. R: A language and environment for statistical computing [online]. R Foundation for

864   Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/ [accessed 1 June

865   2014].

866   Regehr, H.M., and Montevecchi, W.A. 1997. Interactive effects of food shortage and predation on

867   breeding failure of black-legged kittiwakes: indirect effects of fisheries activities and implications for

868   indicator species. Mar. Ecol.: Prog. Ser. **155**: 249–260. doi:10.3354/meps155249.

869   Reid, R.N., Cargnelli, L.M., Griesbach, S.J., Packer, D.B., Johnson, D.L., Zetlin, C.A., Morse, W.W., and

870   Berrien, P.L. 1999. Essential Fish Habitat Document: Atlantic herring, *Clupea harengus*, life history and

871   habitat characteristics. National Oceanic and Atmospheric Administration Technical Memorandum

872   NMFS-NE-126.

873   Rijnsdorp, A.D., Peck, M.A., Engelhard, G.H., Möllmann, C., and Pinnegar, J.K. 2009. Resolving the effect

874   of climate change on fish populations. ICES J. Mar. Sci., **66**(7): 1570–1583. doi:10.1093/icesjms/fsp056.

875   Scopel, L.C., and Diamond, A.W. 2017. The case for lethal control of gulls on seabird colonies. J. Wildl.

876   Manage **81**(4): 572-580. doi: 10.1002/jwmg.21233

877   Scopel, L.C., and Diamond, A.W. Predation and food-weather interactions drive colony collapse in a

878   managed metapopulation of Arctic Terns (*Sterna paradisaea*). Can. J. Zool. In press.

879   Singh, R., Dalton, A., Knox, D., Macintyre, A., and Melvin, G.D. 2016. 2015 Evaluation of Northwest

880   Atlantic Fisheries Organization (DAFO) divisions 4VWX herring. DFO Can. Sci. Advis. Sec. Res. Doc.

881   2016/073.

882   Skern-Mauritzen, M., Ottersen, G., Handegard, N.O., Huse, G., Dingsør, G.E., Stenseth, N.C., and Kjesbu,

883   O.S. 2016. Ecosystem processes are rarely included in tactical fisheries management. Fish and Fisheries

884   **17**: 165–175. doi:10.1111/faf.12111.

885   Smith, P.C., Pettigrew, N.R., Yeats, P., Townsend, D.W., and Han, G. 2012. Regime shift in the Gulf of

886   Maine. Am. Fish. Soc. Symp. **79**:185-203.

887   Stephenson, R.L. 2012. A perspective on advancing ecosystem research for the Gulf of Maine. Am.

888   Fish. Soc. Symp. **79**:409-415.

40

889 Stephenson, R.L., Lane, D.E., Aldous, D.G., and Novak, R. 1993. Management of the 4WX Atlantic

890 herring (*Clupea harengus*) fishery: an analysis of recent events. Can. J. Fish. Aquat. Sci. **50**(12): 2742-

891 2757. doi:10.1139/f93-299.

892 Stephenson, R.L., Power, M.J., Clark, K.J., Melvin, G.D., Fife, F.J., and Paul, S.D. 1998. 1998 Evaluation of

893 4VWX herring. DFO Can. Sci. Advis. Sec. Res. Doc. 98/52.

894 Stephenson, R.L., Power, M.J., Laffan, S.W., and Suthers, I.M. 2015. Tests of larval retention in a tidally

895 energetic environment reveal the complexity of the spatial structure in herring populations. Fisheries

896 Oceanography **24**(6): 553-570. doi:10.1111/fog.12129.

897 Stevenson, D.K., and Scott, M.L. 2005. Atlantic herring, *Clupea harengus*, life history and habitat

898 characteristics. National Oceanic and Atmospheric Administration Technical Memorandum NMFS-NE-

899 192.

900 Suryan, R.M., Irons, D.B., Brown, E.D., Jodice, P.G.R., and Roby, D.D. 2006. Site-specific effects on

901 productivity of an upper trophic-level marine predator: Bottom-up, top-down, and mismatch effects on

902 reproduction in a colonial seabird. Prog. Oceanogr. **68**(2-4): 303–328. doi:10.1016/j.pocean.2006.02.006

903 Sydeman, W.J., Mills, K.L., Santora, J.A., Thompson, S.A., Bertram, D.F., Morgan, K.H., Wells, B.K.,

904 Hipfner, J.M., and Wolf, S.G. 2009. Seabirds and climate in the California Current – a synthesis of

905 change: troubling trends or meandering variability? California Cooperative Oceanic Fisheries

906 Investigations Reports **50**: 82–104. Available from

907 https://137.110.142.7/publications/CR/2009/2009Syd.pdf [accessed 10 January 2017].

908 Thompson, S.A., Sydeman, W.J., Santora, J. A., Black, B. A., Suryan, R.M., Calambokidis, J., Peterson,

909 W.T., and Bograd, S.J. 2012. Linking predators to seasonality of upwelling: Using food web indicators

910 and path analysis to infer trophic connections. Prog. Oceanogr. **101**(1): 106–120.

911 doi:10.1016/j.pocean.2012.02.001.

912 Townsend, D.W., Rebuck, N.D., Thomas, M.A., Karp-Boss, L., and Gettings, R.M. 2010. A changing

913 nutrient regime in the Gulf of Maine. Cont. Shelf Res. **30**:820-832.

914 Transboundary Resource Assessment Committee. 2009. Gulf of Maine-Georges Bank herring stock

915 complex. Status report 2009/04.

916 Travis, J., Coleman, F.C., Auster, P.J., Cury, P.M., Estes, J.A., Orensan, J., Peterson, G.H., Power, M.E.,

917 Steneck, R.S., and Wootton, T.J. 2014. Integrating the invisible fabric of nature into fisheries

918 management. Proc. Natl. Acad. Sci. U. S. A. **111**(2): 581–584. doi:10.1073/pnas.1305853111.

919 Tupper, M.H., Anthony, V.C., Chenoweth, S.B., and MacCluen, H.A. 1998. Biology and assessment of Gulf

920 of Maine Herring Stocks. Gulf of Maine Aquarium, Portland, ME.

921 Wanless, S., Harris, M.P., Redman, P., and Speakman, J.R. 2005. Low energy values of fish as a probable

922 cause of a major seabird breeding failure in the North Sea. Mar. Ecol. Progr. Ser. **294**:1-8.

923 Wootton, R.J. 1998. Ecology of teleost fishes. Kluwer Academic Publishers, Dordecht, The Netherlands.

924 Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., and Smith, G.M. 2009. Mixed effects models and

925 extensions in ecology with R. Springer, New York, NY.
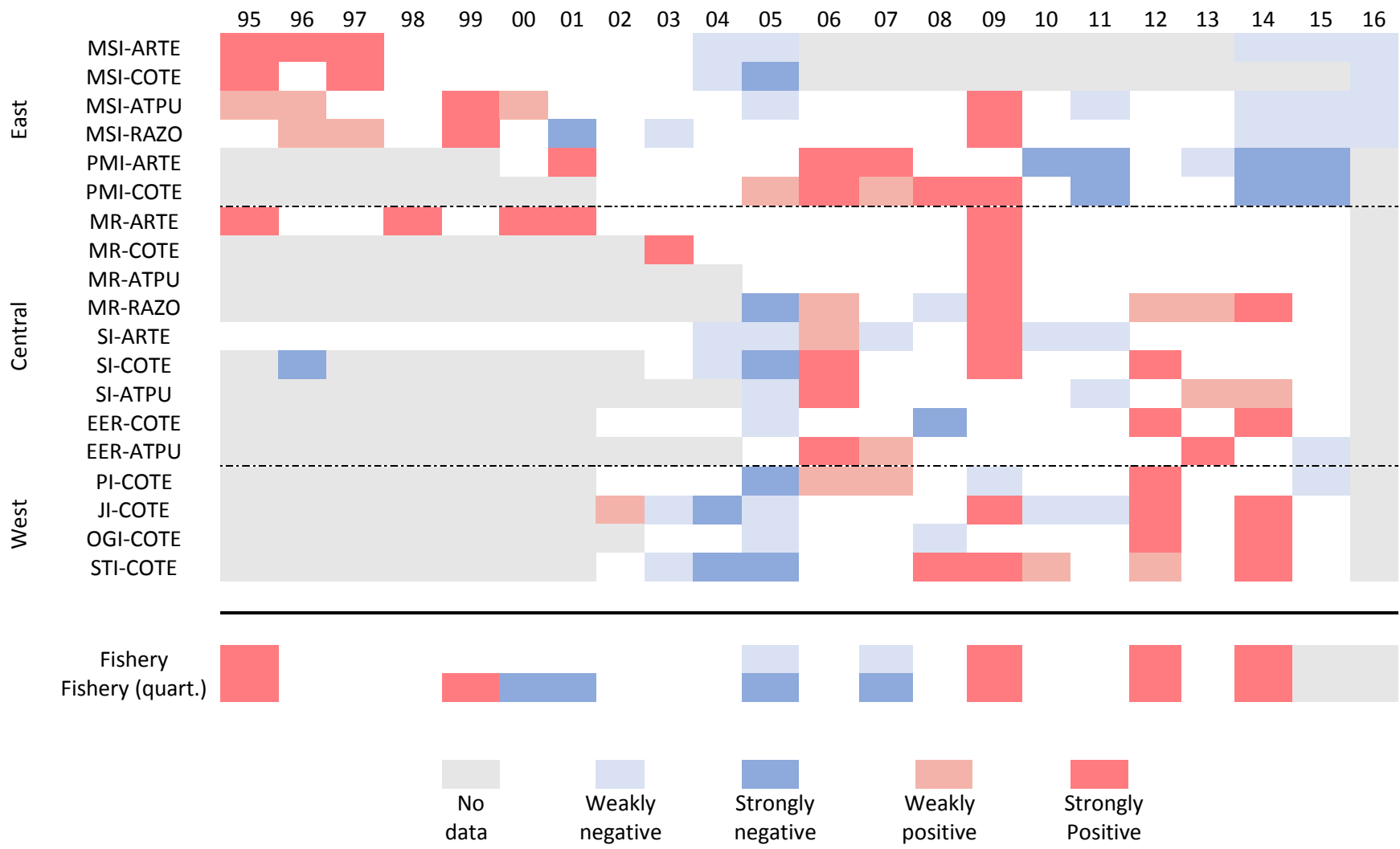
926

927 Table 1. Relative Atlantic herring age-1 cohort strength among seabird colonies and fishery data. Boxes indicate strong (red) or weak (blue)

928 cohorts for a particular seabird colony-species pair, or for cohorts as defined by the American stock assessment model. Fishery cohorts are

929 listed by raw abundance and by quartile. Colonies include Machias Seal Island (MSI), Petit Manan Island (PMI), Matinicus Rock (MR), Seal Island

930 NWR (SI), Eastern Egg Rock (EER), Pond Island NWR (PI), Jenny Island (JI), Outer Green Island (OGI), and Stratton Island (STI). Seabirds are Arctic

931 tern (ARTE), common tern (COTE), Atlantic puffin (ATPU), and razorbill (RAZO).

932   Table 2. List of seabird-fishery predictive models with a prediction error <0.2. Models are ranked by

933   prediction error. All models are considered strong potential bioindicators.

| Fishery Category	Fishery Description	Seabird Colony	Seabird Species	Seabird Data Type	Lag	Prediction Error	SE	Pseudo-R^2	Slope	SE	N
Acoustic	Bay of Fundy	Machias Seal Is.	C. Tern	Proportion	+1	0.03	0.00	1.14	1.00	0.08	8
Model	Spawning Stock	Matinicus Rock	Razorbill	Mass/feed	+5	0.05	0.01	1.00	1.09	0.16	5
Landings	Combined	Machias Seal Is.	C. Tern	Mass/hr	+4	0.08	0.00	0.94	0.64	0.11	11
Landings	Fixed	Machias Seal Is.	C. Tern	Proportion	-2	0.13	0.01	0.86	0.62	0.12	12
Model	Recruits, USA	Matinicus Rock	C. Tern	Mass/hr	0	0.13	0.01	0.87	1.07	0.13	9
Model	Spawning Stock	Jenny Is.	C. Tern	Mass/feed	+5	0.13	0.01	0.91	1.07	0.16	8
Acoustic	Georges Bank	Machias Seal Is.	C. Tern	Mass/hr	+4	0.15	0.01	0.73	1.03	0.14	11
Weight	At-age 1	Seal Is.	Puffin	Avg. mass	+5	0.15	0.02	0.91	0.82	0.24	6
Model	Recruits, USA	Matinicus Rock	Puffin	Proportion	0	0.17	0.02	0.86	1.33	0.22	7
Acoustic	Bay of Fundy	Pond Is.	C. Tern	Proportion	+3	0.18	0.01	0.76	0.97	0.22	9
Landings	Combined	Petit Manan Is.	C. Tern	Proportion	0	0.18	0.01	0.80	0.54	0.15	12
Landings	Mobile, USA	Stratton Is.	C. Tern	Mass/hr	+5	0.18	0.02	0.87	0.64	0.19	8
Acoustic	Georges Bank	Machias Seal Is.	A. Tern	Mass/hr	+4	0.19	0.01	0.65	0.88	0.14	11

934

935

936   Table 3. Rankings of 13 fishery data types for potential use as indices of predator food availability in the

937   Gulf of Maine and Bay of Fundy region. Ranking criteria included the number of seabird species

938   included in cross-validated models (four total), the number of seabird colonies included in cross-

939   validated models (nine total), the number of geographic regions encompassed by the models (three

940   total), the number of cross-validated models with slopes and pseudo-R^2 >0.5, and the number of cross-

941   validated models with prediction error <0.2 represented by each data type.

Category	Description	# Species	# Colonies	# Regions	# Models >0.5	# Error <0.2	Potential?
Model	Spawning Stock Biomass	4	8	3	10	2	High
Acoustic	Bay of Fundy (Canada)	4	5	3	6	2	High
Weight	At-age 1	3	5	2	12	1	Moderate
Landings	Mobile, USA	3	5	2	7	1	Moderate
Model	Recruits, Canada	3	5	2	7	0	Moderate
Landings	Fixed, New Brunswick	3	3	2	6	1	Moderate
Model	Recruits, USA	2	2	1	4	2	Low
Acoustic	Georges Bank (USA)	2	1	1	4	2	Low
Weight	At-age 3-5	3	3	2	4	0	Low
Weight	At-age 3+	2	1	1	4	0	Low
Landings	Combined (USA Mobile, all Fixed)	2	2	1	3	2	Low
Weight	Growth increment 2-3	2	3	2	3	0	Low
Landings	Mobile, New Brunswick	0	0	0	0	0	Poor

942
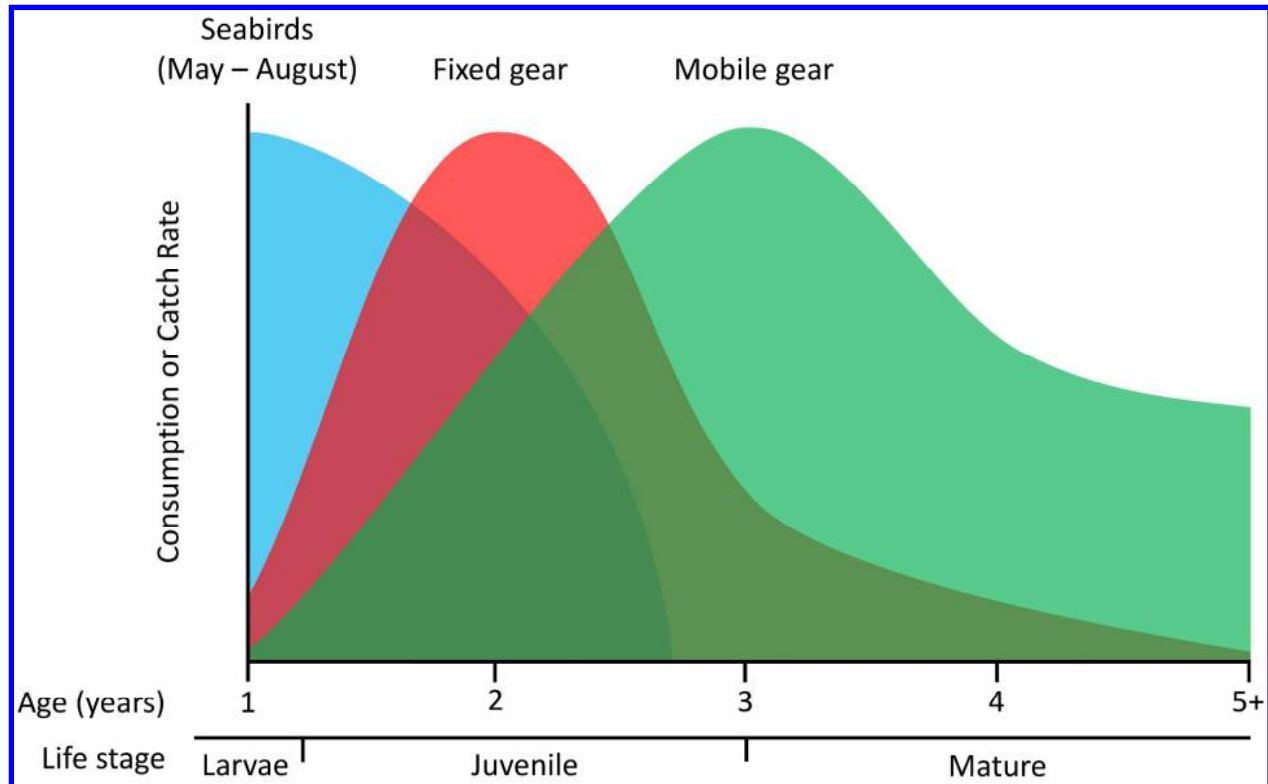
Figure 1. Stages of Atlantic herring life cycle and timing of peak mortality from seabirds and fisheries.
Herring hatch in the fall (age 0, not shown), and become age 1 on January 1; fall-spawned herring do not
overlap with the seabird breeding season at age 0. Seabirds target age-1 herring in summer, and, to a
lesser extent, age-2 herring the following summer. The fixed-gear fishery primarily catches age-2
herring, and the mobile-gear fishery catches a range of fish cohorts, especially ages 3-5.
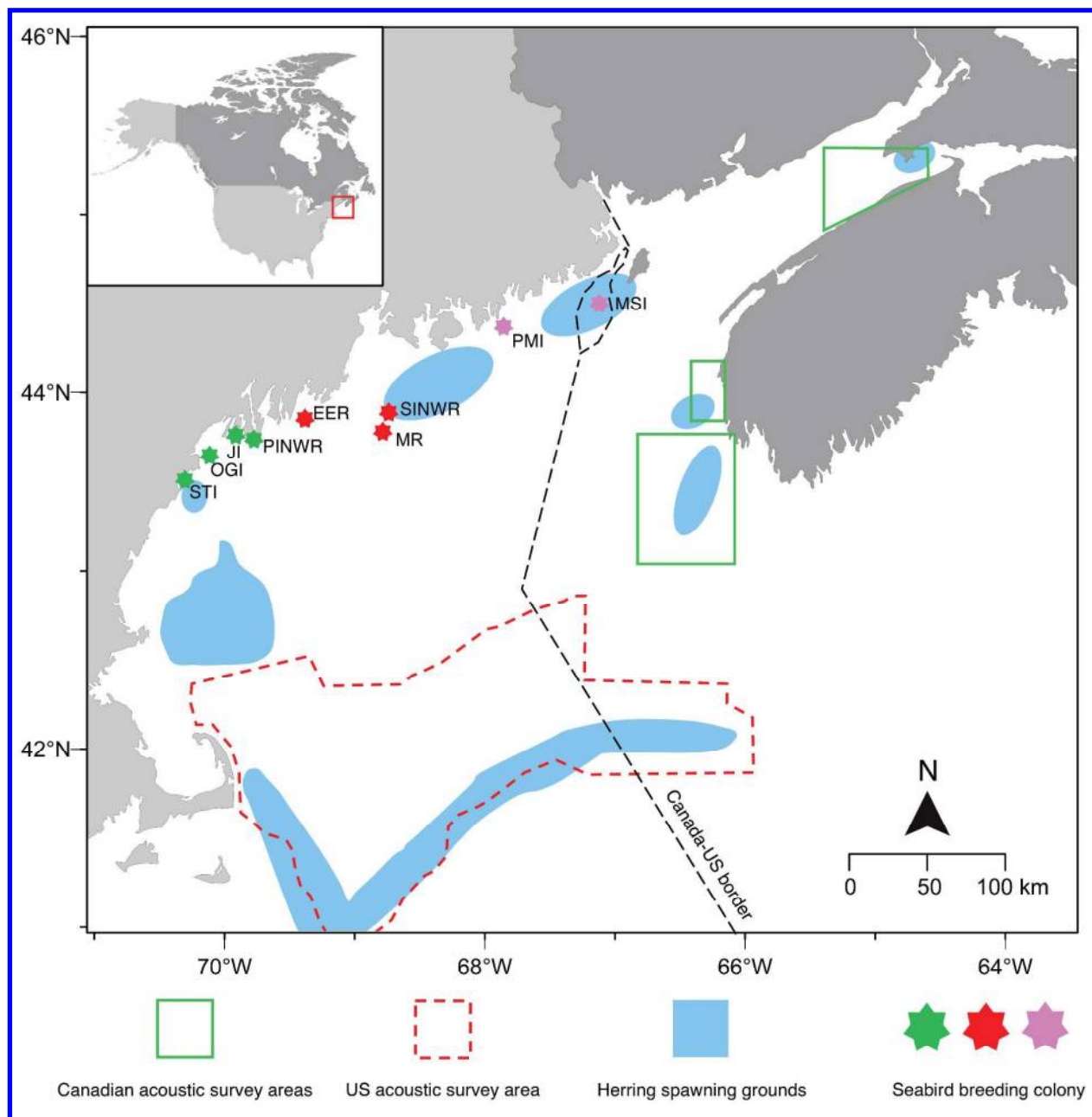
Figure 2. Map depicting the Gulf of Maine-Bay of Fundy region, the nine seabird colonies in this study, spawning grounds of Atlantic herring, and locations of acoustic surveys. Colonies are annotated as follows: Stratton Island (STI), Outer Green Island (OGI), Jenny Island (JI), Pond Island National Wildlife Refuge (PINWR), Eastern Egg Rock (EER), Matinicus Rock (MR), Seal Island National Wildlife Refuge (SINWR), Petit Manan Island (PMI), and Machias Seal Island (MSI). Colonies in green are considered

957   western, colonies in red are considered central, and colonies in pink are considered eastern. Adapted

958   from Overholtz et al. 2004; NEFSC 2012; DFO 2016.
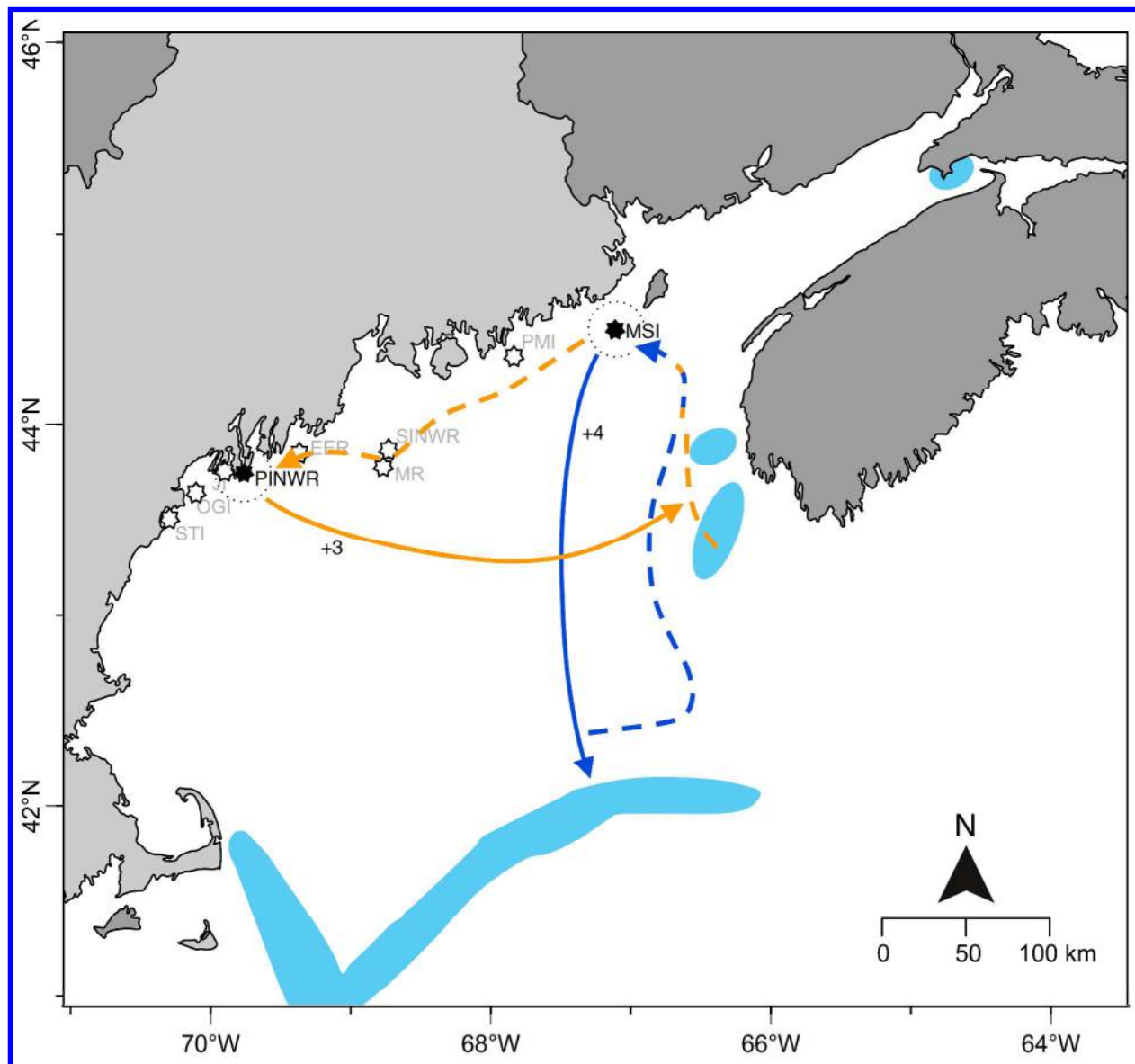
959
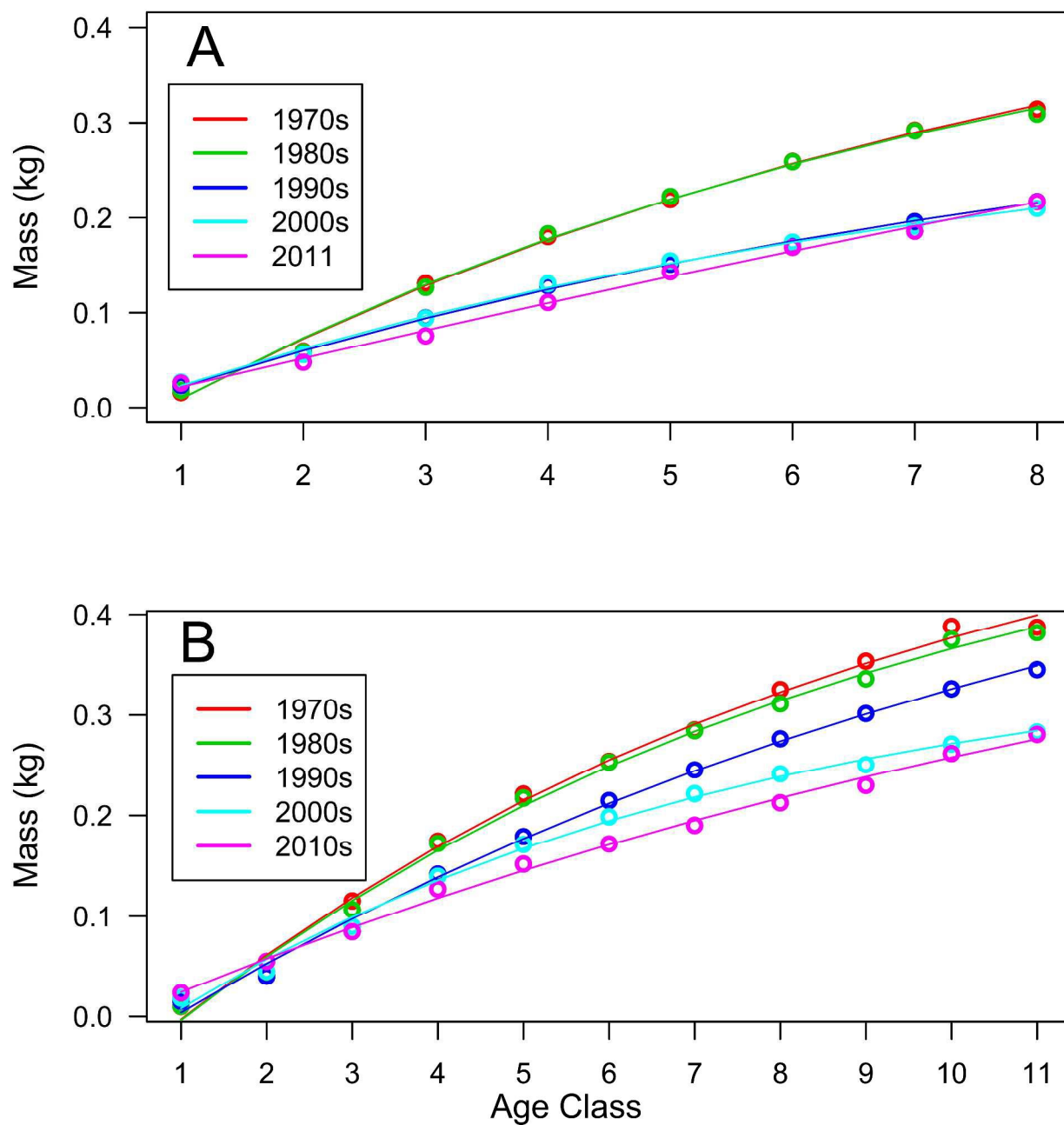
960

961   Figure 3. Mean weights-at-age for the (A) American stock assessment and (B) Canadian stock

962   assessment. American data include 1965-2011, while the Canadian data include 1965-2015.

963

964

965 Figure 4. Top ranked models showing relationships between seabird diet data and American (blue) and

966 Canadian (orange) acoustic surveys in the Gulf of Maine and Bay of Fundy region. Common and Arctic

967 tern diet data at Machias Seal Island (MSI) predicted acoustic surveys at Georges Bank four years later,

968 while common terns at Pond Island (PINWR) predicted acoustic surveys from the Bay of Fundy three

969 years later. Dashed lines suggest proposed larval movements from spawning areas toward seabird

970   colonies, where seabirds consume juvenile herring. Solid lines indicate connectivity back to natal

971   grounds.

972
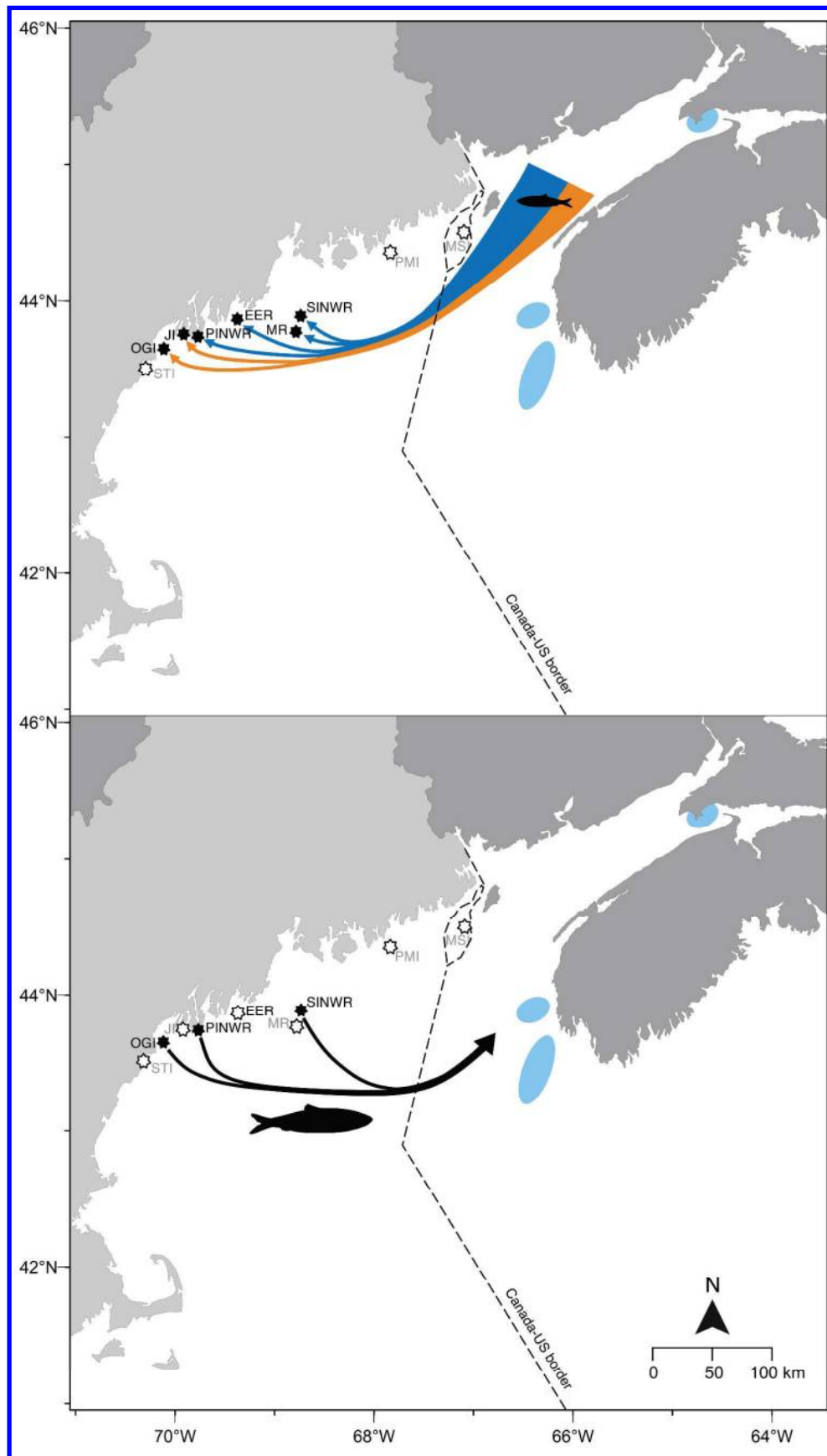
973   Figure 5. Links between Canadian fishery data and seabird diets in the Gulf of Maine and Bay of Fundy

974   region, based on cross-validated regression models. Two types of fishery data represented the age-1

975   cohort (age-1 recruitment, weight-at-age 1, upper panel), and one was linked to older cohorts (acoustic

976   survey, lower panel). The blue and orange arrows indicate links between age-1 data and seabird

977   colonies; the blue arrow indicates a significant relationship with both sets of data, whereas the orange

978   arrow shows a link with just one. The black arrows show ties back to the Bay of Fundy acoustic data

979   three years later, suggesting movement back to Canada three years later. Spawning areas are shown in

980   blue polygons, where acoustic surveys are performed.