# Influence of environmental factors on capelin distributions in the Gulf of Alaska

David W. McGowan[a *], John K. Horne[a].

James T. Thorson[b], Mark Zimmermann[c]

[a] *School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195-5020, USA*

[b] *Fishery Resource Analysis and Monitoring Division (FRAM), Northwest Fisheries Science Center, National Marine Fisheries Service (NMFS), National Oceanic and Atmospheric Administration (NOAA), 2725 Montlake Boulevard East, Seattle, WA 98112, USA*

[c] *Resource Assessment and Conservation Engineering Division (RACE), Alaska Fisheries Science Center, National Marine Fisheries Service (NMFS), National Oceanic and Atmospheric Administration (NOAA), 7600 Sand Point Way NE, Seattle, WA 98115, USA*

[*] *Correspondence*

*E-mail address:* mcgowand@uw.edu (D.W. McGowan)

**ABSTRACT**

Capelin (*Mallotus villosus*) are an important mid-trophic link within marine food webs, yet there is limited information describing fluctuations in capelin distributions and abundances in the Alaskan North Pacific. The influence of physical and biological environmental factors on capelin distributions was investigated over continental shelf waters in the central Gulf of Alaska (CGOA). Acoustic, oceanographic, and trawl sampling were conducted in summer and fall of 2011 and 2013 as part of the Gulf of Alaska Integrated Ecosystem Research Program. Continuous acoustic sampling across bathymetric gradients showed capelin distributions on the CGOA shelf were primarily concentrated over submarine banks less than 100 m in depth. Acoustic densities of capelin were higher over banks in summer of both years compared to in deeper troughs (≥ 100 m depth). Length-based distributions indicated that capelin occurring over banks were mostly age-1 fish, while larger age-2+ capelin were more prevalent in troughs. Environmental factors that influenced capelin occurrence and density in summer 2013 were identified at two spatial resolutions associated with systematic sampling at discrete stations and continuous sampling along transects. At the station resolution (37 km), the probability of capelin occurrence was associated with relatively warmer bottom temperatures and increased chlorophyll concentrations, while capelin density was associated with close proximity to the edges of banks and reduced thermal stratification. At the transect resolution (0.5 km), increases in capelin density were associated with increases in density of a potential competitor (age-0 pollock, *Gadus chalcogrammus*) over banks and with increases in prey (macrozooplankton) and predator (semi-demersal groundfish) densities in troughs. We infer that in summer, age-1 capelin concentrate over shallow banks on the CGOA shelf in waters that are well-mixed and more productive to feed on abundant copepods, whereas age-2+ capelin primarily occupy deeper waters in troughs where they

41 consume larger euphausiid and amphipod prey. This suggests that distributions of both life stages over
42 the shelf were likely driven by changes in the availability and composition of prey.
43
44 **1. INTRODUCTION**
45
46 In northern, high-latitude ecosystems, capelin (*Mallotus villosus*) are an important mid-trophic
47 link within marine food webs. A mobile and gregarious small pelagic species, capelin are spatially and
48 temporally heterogeneous (Gjøsæter, 1998; Rose, 2005; Vilhjálmsson, 2002), and function as a predator
49 of zooplankton and prey for piscivorous fish, seabirds, and marine mammals. As a result, fluctuations in
50 capelin distributions and abundances modulate the transfer of energy from lower to upper trophic levels
51 in marine food webs throughout boreal waters (Gjøsæter et al., 2009; Hjermann et al., 2010;
52 Vilhjálmsson, 2002).
53 Understanding how capelin biology, population dynamics, and distributions are influenced by
54 environmental processes has been a research focus spanning decades in the Atlantic. Fluctuations in
55 capelin distributions and abundances are well documented in the Barents Sea, Icelandic, and Northwest
56 Atlantic populations (Carscadden et al., 2013a, 2013b; Gjøsæter, 1998; Vilhjálmsson, 2002), with
57 fisheries and oceanographic data to support commercial capelin fisheries (e.g. Carscadden et al., 2013a;
58 Gjøsæter, 1998). These data have facilitated investigations on the potential role of directed fisheries on
59 population crashes, and how environmental processes influence capelin population dynamics and
60 distributions during different life stages (e.g. Carscadden et al., 2013b; Frank et al., 1996; Ingvaldsen and
61 Gjøsæter, 2013; Olafsdottir and Rose, 2012). Studies on Atlantic capelin populations also highlight how
62 the strength and direction of capelin-predator relationships, including Atlantic cod (*Gadus morhua*)
63 (Horne and Schneider, 1994; Piatt, 1990; Rose and Leggett, 1990) and seabirds (Fauchald, 2009;
64 Fauchald and Erikstad, 2002), vary as a function of scale.

Compared to Atlantic populations, there is limited information describing changes in Pacific capelin distributions and abundances. This lack of information is attributed to the absence of a directed capelin fishery, and that existing abundance surveys for demersal fish species are poorly suited to detect changes in capelin and other small pelagic species' distributions (Ormseth, 2012). Despite limited monitoring, interdecadal abundance fluctuations have been observed in the GOA (Anderson and Piatt, 1999; Ormseth, 2012), and are believed to have impacted predators. After the late 1970s regime shift (Francis et al., 1998), declines in production and abundances of piscivorous seabirds and marine mammals were attributed to reductions of capelin and other forage species in predator diets (Anderson and Piatt, 1999; Merrick et al., 1997; Piatt and Anderson, 1996). Recognition that changes in the availability and species composition of prey may impact managed predators (Ormseth, 2012; Springer and Speckman, 1997) prompted investigations of how physical and biological processes influence the distribution and abundance of capelin and other small pelagic species in the Northeast Pacific (e.g. Arimitsu et al., 2008; Hollowed et al., 2012; Parker-Stetter et al., 2016).

Recent interdisciplinary research efforts in Alaska have attributed a wide range of environmental factors to changes in offshore and coastal capelin distributions. In the Gulf of Alaska (GOA), changes in capelin distributions over the continental shelf east of Kodiak Island have been attributed to oceanographic conditions and the location of water masses (Hollowed et al., 2007; Logerwell et al., 2007), composition and distribution of zooplankton prey, and potential competition with juvenile walleye pollock (*Gadus chalcogrammus*, hereafter pollock) (Logerwell et al., 2007, 2010; Wilson et al., 2006). GOA-based analyses have also shown differences in capelin diets within and between years that were associated with capelin length and water mass type (Logerwell et al., 2010; Wilson et al., 2006). Within Glacier Bay in southeast Alaska, capelin have been associated with areas near tidewater glaciers (Arimitsu et al., 2008). In the eastern Bering Sea (EBS), the relative abundance of capelin in surface waters (upper 30 m) was significantly higher during years with colder water (2006-

2011), during which distributions were concentrated in the northeastern Bering Sea and extended south compared to observations from warm years (2003-2005) (Andrews et al., 2015). Capelin located in surface waters during cold years were predicted to occur in relatively cooler, less saline waters above the pycnocline where densities would be highest over bottom depths between 60 to 80 m (Parker-Stetter et al., 2016).

This study was conducted to identify physical and biological factors that influence capelin distributions in the GOA over the shelf east of Kodiak Island. Our investigation was facilitated by a fisheries-oceanographic survey that measured distributions of pelagic fishes and macrozooplankton. As part of the Gulf of Alaska Integrated Ecosystem Research Program (GOAIERP, http://www.nprb.org/gulf-of-alaska-project), acoustic-trawl surveys were conducted in summer and fall 2011 and summer 2013. GOAIERP was an interdisciplinary effort to investigate how recruitment dynamics of five key groundfish species are influenced by physical and biological processes during their first year of life (Moss et al., 2016b). Recognizing that the relative influence of explanatory factors on capelin distributions may be scale-dependent (Schneider, 1994), we analyzed data at station (37 km) and transect (0.5 km) spatial resolutions. Our specific objectives were to: 1) quantify the influence of physical and biological factors on capelin occurrence and density; and 2) compare the relative importance of factors that explain variance in capelin distributions at two spatial resolutions.

## 2. METHODS

### *2.1. Study area and survey design*

Data were collected in the central GOA (CGOA) during GOAIERP acoustic-trawl surveys on the F/V *Northwest Explorer* in summer (5-21 August) and fall (25 September to 8 October) 2011, and summer (6-21 August) 2013 (see McGowan et al., 2016 for details). The study area extended from the southeast side of Kodiak Island to the mouth of Amatuli Trough (Fig. 1). Acoustic measurements were sampled along parallel transects orthogonal from the coast (< 50 m bottom depth) to basin waters beyond the 2000 m depth contour. Midwater trawls were conducted opportunistically along transects to sample observed acoustic targets. Surface trawl and oceanographic measurements were sampled at fixed stations spaced equidistant at 37 km (20 nm) along 10 transects that ranged from 111 to 185 km in length (4 to 6 stations per transect) depending on the width of the shelf. All sampling was conducted during daytime hours (i.e. 30 minutes after sunrise to 30 minutes before sunset).

*2.2. Data collection and processing*

*2.2.1. Acoustic and trawl data*

Acoustic data were collected using a Simrad ES-60 echosounder (Kongsberg Maritime) with hull-mounted, split-beam transducers operating at 38 and 120 (added in 2013) kHz. Prior to the start of summer surveys, both frequencies were calibrated using a 38.1 mm tungsten carbide sphere following Foote et al. (1987). Fish were sampled at surface (0-30 m) and midwater depths (30-250 m) using a 198 m midwater rope trawl (Cantrawl model 400, Cantrawl Nets Ltd) with a 1.2 mm mesh cod-end liner. Trawl samples were used to validate the species and length compositions of observed acoustic patterns to partition backscatter (i.e. reflected sound) to single- or mixed-species classification categories: capelin, piscivorous fish (i.e. groundfish species > 158 mm), age-0 pollock, macrozooplankton (e.g. euphausiids and amphipods), capelin/age-0 pollock mix (hereafter forage fish mix), capelin/age-0 pollock/piscivores mix (hereafter forage fish/piscivores mix), and unknown following McGowan et al.

(2016). Backscatter observed during trawls in which one classification category accounted for at least 90% of the catch (by number) was treated as a characteristic echogram pattern for subsequent visual classification of that category. The forage fish mix and forage fish/piscivores mix categories were comprised of backscatter that could not be assigned to capelin, age-0 pollock, or piscivorous fish categories; this occurred when trawl samples indicated that echogram patterns were comprised of multiple forage fish species, or when echogram patterns of two or more categories overlapped. Backscatter was only assigned to one classification category, and patterns that were not sampled by trawls were classified as unknown and excluded from analysis. Categorized backscatter for capelin and piscivorous fish were available for all CGOA surveys. Backscatter data for age-0 pollock and macrozooplankton were only available in summer 2013 due to the rare occurrence of age-0 pollock and lack of 120 kHz data needed to acoustically discriminate macrozooplankton from fish in 2011 (*cf.* McGowan et al., 2016). Along-transect estimates of acoustic density (nautical area scattering coefficient (NASC), $s_A$, $m^2\ nm^{-2}$, MacLennan et al., 2002) for each classification category were calculated by integrating backscatter through the water column from 10 m below the surface to a maximum depth of 250 m or 1 m above bottom. The 10-m upper limit accounted for transducer depth (4.9 m) and excluded data from within twice the near-field of the 38 kHz transducer (Simmonds and MacLennan, 2005). Data below 250 m were excluded due to a lack of trawl samples for species validation. Data within 1 m of the bottom were excluded to account for the acoustic deadzone (Ona and Mitson, 1996). Acoustic density estimates were integrated and exported in 0.2 km horizontal bins, the minimum resolution of acoustic data. Each 0.2 km bin contained a minimum of 5 pings when the echosounder was operated at the slowest pulse rate of 1 ping per 7 seconds when sampling was conducted over deep water (>4000 m). A previously conducted (McGowan et al., 2016) wavelet analysis (Bradshaw and Spies, 1992; Torrence and Compo, 1998) of the 0.2 km GOAIERP acoustic data determined that a 0.5 km horizontal resolution maximized variability in capelin distributions and was used in further analyses.

Additional details of acoustic data collection, processing, and classification can be found in McGowan et al. (2016).

*2.2.2. Oceanographic data*

Oceanographic data were collected at each station. A SeaBird SBE 911 plus (Seabird Electronics) was used to measure conductivity, temperature, and depth (CTD) on the downcast to 200 m or 10 above bottom, whichever was shallower. CTD measurements were error-checked and integrated in 1 m depth intervals at the University of Alaska Fairbanks in Fairbanks, AK (2011 data) and the Pacific Marine Environmental Laboratory (PMEL) in Seattle, WA (2013 data) (*cf.* Stabeno et al., 2015). *In situ* chlorophyll-*a* concentrations were measured from water samples collected on the CTD upcast at 10 m intervals from 50 m depth to the surface. Trapezoidal integration was used to approximate integrated chlorophyll from 0-50 m (*cf.* Strom et al., 2015).

*2.2.3. Bathymetric data*

To investigate the affiliation of capelin with bathymetric features, high resolution bathymetry data (Zimmermann and Prescott, 2015) for the CGOA shelf were used to determine the proximity of acoustic samples to the edges of shallow, submarine banks by measuring the shortest distance from each sample's midpoint to the 100 m bottom depth contour (hereafter *Edge*) in ArcGIS v10.3 (Environmental Systems Research Institute, Inc). Enhanced vertical mixing and primary production occur near the edges of Albatross and Portlock Banks, primarily due to strong tidal mixing that supplies cold, nutrient-rich bottom water from adjacent troughs to the euphotic zone over banks throughout summer (Cheng et al., 2012; Mordy et al., 2016). This variable was added as it was observed that capelin densities were highest over waters less than 100 m bottom depth in summer of both years (McGowan et al., 2016), and appeared to concentrate near the edges of Portlock and Albatross Banks.

184

185 2.3. *Response and predictor indices*

186

187 Response variables occurrence (*P*) and density (*D*) were used to evaluate factors influencing
188 capelin distributions (Table 1). Capelin occurrence was a binary variable (present: *P*=1; absent: *P*=0)
189 derived from acoustic density estimates ($s_A$) for capelin, forage fish mix, and forage fish/piscivores mix
190 classification categories (see Table 2 in McGowan et al., 2016). Capelin were considered present in an
191 acoustic sample when the $s_A$ value was greater than 0 for any of the three categories. Capelin density
192 (*D*) was a continuous measure of nonzero acoustic density (i.e. where fish were present, capelin $s_A > 0$).
193 Using categorical and continuous response variables accounts for limitations in survey design, where the
194 frequency of zero density values in a sample is sensitive to changes in the extent of the survey domain
195 (Fig. 1), and acoustic classification (i.e. backscatter assigned to multi-species or unknown classification
196 categories). Density estimates for forage fish mix and forage fish/piscivores mix categories were not
197 included in analyses of capelin nonzero densities because backscatter could not be apportioned to
198 species from trawl catches due to an inability to correct for net selectivity (e.g. De Robertis et al., 2017a;
199 Williams et al., 2011). Data for all response and predictor indices were limited to observations in waters
200 shallower than 500 m depth over the CGOA shelf or adjacent to the shelf break.
201 Predictor indices used in this study represent the environment where capelin occur, while
202 minimizing collinearity among covariates (Table 1). Bathymetric features were represented by acoustic-
203 based bottom depth (*BtmD*) measurements and the proximity of acoustic samples to the nearest edge
204 of banks (*Edge*) derived from the Zimmermann and Prescott (2015) data. Oceanographic predictor
205 indices included: surface temperature (*TmpS*), measured 1 m below surface; bottom temperature
206 (*TmpB*) and salinity (*SalB*), measured at 200 m or approximately 10 m above bottom; and water column
207 stratification represented by the difference between surface and bottom temperature (*TmpD*). Primary

production was characterized using integrated chlorophyll-*a* (*Chla*) from 0 to 50 m. Acoustic-based, biological predictor indices included: predator density (*Pred*) using the piscivorous fish classification category; a competitor (*Comp*) represented by the age-0 pollock category; and prey density (*Prey*) using data from the macrozooplankton category. *Chla*, *Pred*, *Comp*, and *Prey* were log-transformed prior to analysis to linearize their relationship with response variables (Zuur et al., 2009).

*2.4. Analysis resolution*

The influence of physical and biological factors on capelin distributions over the CGOA shelf were analyzed at two spatial resolutions (hereafter analysis resolution) to differentiate continuous acoustic sampling along transects (hereafter transect-based) from discrete sampling at fixed stations (hereafter station-based). Transect-based analyses used continuous data at a spatial resolution of 0.5 km between midpoints of adjacent samples ($n$=2320 samples in summer 2011; $n$=1636 in fall 2011; $n$=2268 in summer 2013). Station-based analyses examined data at a spatial resolution of 37 km between samples; using both oceanographic measurements from stations ($n$=32 stations sampled in summer 2011; $n$=28 in fall 2011; $n$=37 in summer 2013) and a subset of acoustic samples that were converted to discrete measurements of mean acoustic density at each station. Mean acoustic density was calculated at each station by averaging all 0.2 km resolution samples located within 0.5 km from a station's CTD cast location and collected the same day as the CTD cast. For stations with less than 3 acoustic samples within 0.5 km of the CTD cast due to logistical constraints, the closest 3 acoustic samples were used. The small number of acoustic samples used to estimate mean acoustic density at each station (3 to 6, 0.2 km samples) was intended to represent the area swept by a typical pelagic (e.g. Wilson, 2009) or surface (e.g. Farley et al., 2005) trawl sample at each station.

Oceanographic predictor indices were not included in transect-based analyses due to the concern that station-based values from field samples interpolated from a 37-km sample resolution to a

232 0.5-km analytic resolution would not represent oceanographic conditions experienced by capelin during
233 acoustic sampling. Heterogeneity in patterns associated with water masses, primary production, and
234 zooplankton communities over the CGOA shelf (e.g. Coyle et al., 2013; Hopcroft et al., 2016; Waite and
235 Mueter, 2013) can occur at scales less than the spatial (37-km) and temporal (approximately 4 hours to
236 4 days) sampling resolution of station-based data. Since the goal of this study is to examine the
237 influence of physical and biological factors on capelin distributions using assumed instantaneous field
238 observations of capelin and their environment, oceanographic indices were only included in station-
239 based analyses.

240 To characterize environmental conditions where capelin occurred, all predictor indices are
241 summarized relative to the presence or absence of capelin. Capelin densities were also plotted relative
242 to each oceanographic index. For each biological predictor index, we qualitatively assessed similarity
243 between patterns at transect- and station-based resolutions by comparing the relative difference in
244 summarized predictor values (i.e. median values, range between $25^{th}$ and $75^{th}$ quantiles) when capelin
245 were present versus absent.

246

247 *2.5. Modeling approach*

248

249 *2.5.1. Model structure*

250 A spatial generalized linear mixed effects model (spatial GLMM) adapted from Thorson et al.
251 (2015a) was used to predict capelin occurrence and density using fixed effects for predictor indices
252 (Table 1), and random effects for spatial covariance at each sample location. Treating spatial covariance
253 as a random effect enables a stochastic process to represent the cumulative impact of physical and
254 biological factors on capelin distributions that are not directly measured (Dormann et al., 2007; Thorson
255 et al., 2015a).

Probability of occurrence, $\hat{p}_i$, for sample $i$ is estimated as

$$\Pr[P=1] = \hat{p}_i = \left(1 - e^{-e^{\lambda_i^{(p)}}}\right) \quad (1)$$

where $P$ is fitted to a Bernoulli distribution with a complementary log-log link (i.e. clog-log) (Hardin and Hilbe, 2007). The asymmetrical shape of the clog-log link function is better-suited to model binary responses with a disproportionately greater number of zeros relative to ones (Hardin and Hilbe, 2007). The linear predictor, $\lambda_i^{(p)}$, for sample $i$ combines fixed and random effects as

$$\lambda_i^{(p)} = \sum_{k=1}^{n_k} x_{i,k}\hat{\beta}_k^{(p)} + \hat{\omega}^{(p)}(s_i) \quad (2)$$

where $x_{i,k}$ is covariate $k$ for sample $i$, $\hat{\beta}_k^{(p)}$ is the parameter for covariate $k$ estimated as a fixed effect, and $\hat{\omega}^{(p)}(s_i)$ is the spatial residual at location $s_i$ estimated as a random effect. Spatial covariances are represented by a Gaussian Markov random field as

$$\hat{\omega}^{(p)} \sim MVN\left(0, \Sigma_\omega^{(p)}\right) \quad (3)$$

where $MVN$ is a multivariate normal distribution fixed at 0, and $\Sigma_\omega^{(p)}$ is the covariance of the random field at each sample location (i.e. spatial residuals) following a Matérn distribution with smoothness $\nu$ = 1 (Thorson et al., 2015a, 2015b). The properties of the random are represented by the marginal standard deviation and the geostatistical range; the former indicates the standard deviation at any location when simulating a new realization from the random field and the latter indicates the distance at which correlations decline to 10% (*cf.* Thorson et al., 2015a) .

Capelin nonzero density, $d_j$, for sample $j$ (where capelin $s_A > 0$) is estimated as

$$\Pr[D = d \mid D > 0] = \text{Gamma}(d_j, \sigma^{-2}, \exp\left(\lambda_j^{(d)}\right)\sigma^2) \quad (4)$$

where $D$ is fitted to a gamma distribution with shape $\sigma^{-2}$ and scale $\exp\left(\lambda_j^{(d)}\right)\sigma^2$, and $\sigma$ is the coefficient of variation for measurement error. The linear combination of fixed and spatial random effects, $\lambda_j^{(d)}$, for sample $j$, is approximated with a log link as

$$\lambda_j^{(d)} = \sum_{k=1}^{n_k} x_{j,k}\hat{\beta}_k^{(d)} + \hat{\omega}^{(d)}(s_j) \quad (5)$$

where $x_{j,k}$ is covariate $k$ for sample $j$, $\hat{\beta}_k^{(d)}$ is the estimated parameter for covariate $k$, and $\hat{\omega}^{(d)}(s_j)$ is the estimated spatial residual at location $s_j$.

Following Thorson et al. (2015b), we used the stochastic partial differential equation (SPDE) approximation (Lindgren et al., 2011) to estimate the conditional probability of the spatial covariance matrix, $\Sigma_\omega$. The three components of the precision matrix used in the SPDE approximation (*cf.* Lindgren et al., 2011) were derived using the R package *INLA* (http://www.r-inla.org, Rue et al., 2009). Template Model Builder (TMB, Kristensen et al., 2016) was used to estimate all parameters using maximum marginal likelihood. TMB calculates the marginal likelihood of fixed effects when integrated across random effects using the Laplace approximation (Skaug and Fournier, 2006). The marginal likelihood was then optimized using a conventional nonlinear optimizer in the R statistical environment (http://www.R-project.org, R Core Development Team, 2013) to estimate fixed effects by maximizing the log-marginal likelihood (Thorson et al., 2015a, 2015b).

*2.5.2. Data analyses and model selection*

To quantify the influence of physical and biological factors on capelin distributions, a multistep approach was used to first quantify changes in nonzero density (i.e. where fish were present) relative to bottom depth by year and season, and then to identify which predictor indices (Table 1) best explained variance in capelin occurrence and density in summer 2013. The first step was based on McGowan et al.'s (2016) preliminary analysis of capelin distributions in summer and fall of 2011 and summer 2013

298 that showed average densities between major depth contours were higher at depths less than 100 m
299 bottom depth relative to densities over deeper waters in summer of both years, and that overall
300 densities were higher in 2013 compared to 2011. To assess if the spatial GLMM would estimate similar
301 depth-related differences in density between seasons and years across a continuous bathymetric
302 gradient, capelin nonzero densities were quantified relative to bottom depth (*BtmD*), season (*fSn)*, and
303 year (*fYr*) using continuous transect data (i.e. 0.5 km resolution). The linear predictor, $\lambda_j^{(d)}$, from Eq. (5)
304 included interaction terms for *BtmD*fSn* and *BtmD*fYr* to allow slopes to vary between surveys.
305 Because capelin distributions were assumed independent among surveys, spatial effects, $\omega_z$, estimated
306 in Eq. (3) required a separate covariance matrix, $\Sigma_{\omega,z}$, for each survey, $z$.

307 In addition to examining differences in density relative to bottom depth, year, and season,
308 length distributions of capelin over banks and troughs were also compared among surveys. Length
309 measurements from midwater and surface trawl catches were pooled by survey and bottom depth
310 factor (*fBT*, Table 1).

311 All subsequent analyses that assess the influence of physical and biological factors on capelin
312 distributions were limited to data from the summer 2013 survey due to the relatively high proportion of
313 acoustic samples located over the CGOA shelf classified as unknown backscatter in 2011 (Fig. 1), and the
314 availability of age-0 pollock and macrozooplankton acoustic data *(cf.* McGowan et al., 2016). The high
315 proportion of unknown backscatter in 2011 resulted from a low number of trawl samples used to verify
316 observed acoustic targets, which would potentially confound estimates of capelin occurrence due to the
317 inability to differentiate between true and false (i.e. sample) zeroes of capelin density.

318 To identify which physical and biological factors explained variance in capelin distributions,
319 model selection was used to identify the best combination of fixed and random effects that predict
320 occurrence and density at station- and transect-based analysis resolutions. All predictor indices were
321 organized into candidate models for each response variable and independent covariates at both analysis

resolutions (Table 1). Each candidate model's name indicates its analysis resolution ("s"=station, "t"=transect), response variable ("P"=presence/absence, "D"=density), and sequential number for each combination of independent covariates. Collinearity among potential predictors was assessed using pairwise correlations (Pearson correlation coefficients, $r$) and variance inflation factors (VIF) (Zuur et al., 2010, 2009). Candidate models were initially organized using pairwise comparisons, in which covariate pairs were considered to be independent if the magnitude of $r$ was less than $|0.5|$ (Zuur et al., 2009). To assess multicollinearity among two or more predictors, VIFs were calculated for each candidate model (Kutner et al., 2004). If one or more predictors had a VIF greater than 2 (Burnham and Anderson, 2002), then the covariate with the highest VIF was dropped. VIFs for the remaining covariates were recalculated, and the process was repeated until all remaining covariates had VIF values less than 2 (Zuur et al., 2010). All predictor indices were assigned to at least one candidate model for analyses of occurrence and density at each analysis resolution. To allow direct comparison of parameter estimates among candidate models for the same response variable and analysis resolution, predictor indices were standardized (i.e. subtract mean and divide by standard deviation) prior to analysis.

Model selection for fixed effects used the stepwise backward approach and Akaike information criterion values corrected for finite sample size (AICc) (Burnham and Anderson, 2002). For each candidate model, the full model was run and AICc calculated using the R package *TMBhelper* (https://github.com/kaskr). The predictor with the highest Wald Z-test p-value was dropped, the model was re-run, and the AICc value recalculated. The reduced model AICc was subtracted from the full model AICc (hereafter ΔAICc). If ΔAICc was greater than 2, then the process was repeated until ΔAICc was less than 2 from the previous step, or all potential predictors had been dropped. The reduced model with a ΔAICc value less than 2 from the previous step is referred to as the final model. Model selection for random effects was based on the relationship of the parameter estimate and its standard error. If the parameter estimate was within one standard error of having no effect, then we inferred

346 that the random effect explained little variance in the model and could be dropped. Sign deviance
347 residuals (*cf.* Hardin and Hilbe, 2007) from models in which spatial effects were dropped were plotted
348 against latitude and longitude to confirm that residuals did not contain any additional spatial structure
349 (Zuur et al., 2009, 2010).

350 For each response variable and analysis resolution, the best model was determined by: the
351 lowest AICc value from the final model of each candidate model; visual inspection of model fit; and
352 proportion of explained deviance (i.e. pseudo $R^2$) (Zuur et al., 2012, 2009). To visually assess model fit,
353 we compared predicted to observed occurrence probabilities using a diagnostic function adapted from
354 the R package *SpatialDeltaGLMM* (https://github.com/James-Thorson). Observed occurrence
355 probabilities were calculated from binary data by sorting predicted probabilities in 0.1 increments and
356 then averaging the corresponding observed binary responses within each of the 0.1 bins (i.e. the
357 average of observed 0s and 1s within each 0.1 increment of predicted occurrence probabilities). For
358 density models, standard diagnostic plots (e.g. residual scatter plots and Q-Q plots) were used to assess
359 model fits (Zuur et al., 2009).

360 To examine the relative importance of physical and biological factors that explained variance in
361 capelin distributions at different analysis resolutions, the best models for transect-based analyses were
362 compared to station-based models comprised of only continuous predictors. Comparisons of parameter
363 estimates between response variables and/or analysis resolutions were qualitatively evaluated using:
364 predictors retained in final reduced models at each resolution; signs of parameter estimates; and
365 magnitudes of parameter estimates relative to other parameters in the same model. To facilitate direct
366 comparison of parameter estimates among continuous covariates within each model, each predictor
367 index was standardized by subtracting values from its mean and dividing by one standard deviation prior
368 to running the model.

369

## 3. RESULTS

Capelin were present in approximately 23% of 6224 acoustic samples (0.5 km data resolution) collected over the CGOA shelf during GOAIERP surveys in 2011 and 2013 (Fig. 1). Among the 97 stations sampled, capelin were present at 27.8%, including 6 of 32 stations sampled in summer 2011, 9 of 28 in fall 2011 (including 3 based on acoustic data classified as mixed forage fish species), and 12 of 37 in summer 2013. The absence of capelin could not be confirmed at 70% of the stations sampled in 2011 due to the presence of acoustic samples with backscatter classified as unknown. All stations sampled in 2013 were included in the analysis as the presence or absence of capelin could be determined for all acoustic samples located within 0.5 km of each station (i.e. no backscatter associated with these samples were classified as unknown).

### *3.1. Distribution relative to bottom depth, season, and year*

Model estimates indicated that capelin nonzero densities were higher in 2013 than in 2011 (Fig. 2A-C). Differences in density relative to bottom depth were observed between seasons, and to a lesser extent between years, as indicated by model interaction terms (Table 2). In summers of both years, densities declined from shallow to deeper waters over the CGOA shelf, although at a slower rate in 2013 as indicated by the coefficient value for the bottom depth (*BtmD*) and year (*fYr*) interaction term ($\hat{\beta}_{BtmD*fYr:2013}$=0.12). In contrast, densities in fall 2011 were lower over shallow waters compared to summer observations, and increased over deeper bottom depths. The spatial GLMM explained approximately 83% of the variation in capelin densities. Diagnostic plots indicate that densities are

underestimated, primarily at lower values, but more than 96% of standardized, sign deviance residuals varied within ± 2 units and lacked trends when compared to predicted densities (Fig. 2D-F).

Capelin distributions in both 2011 and 2013 showed that the 100-m depth contour was an informative delineation. There was a pronounced shift from relatively high to low densities over depths greater than 100 m in summer of both years (2011: mean log $\widehat{D}_{<100\ \mathrm{m}}$=4.7 ± 0.3 $m^2\ nm^{-2}$, mean log $\widehat{D}_{\geq 100\ \mathrm{m}}$=3.1 ± 0.4; 2013: mean log $\widehat{D}_{<100\ \mathrm{m}}$=6.9 ± 0.4 $m^2\ nm^{-2}$, mean log $\widehat{D}_{\geq 100\ \mathrm{m}}$=4.6 ± 0.3) (Fig. 2A-C). Acoustic measurements also showed that capelin were concentrated over shallow waters between 50 to 100 m in all surveys (Fig. 1-2). Based on these patterns, trawl data were partitioned at the 100-m depth contour to examine differences in capelin length frequencies between banks (<100 m) and troughs (≥100 m over the CGOA shelf). Differences in fork length distributions between banks and troughs were evident in all surveys (Fig. 3), with modes and mean lengths smaller over banks. Using Brown's (2002) age-length relationship, observed length modes indicate that banks were primarily occupied by age-1 capelin while age-2+ fish were more prevalent in deeper, trough waters.

### *3.2. Relative importance of physical and biological factors*

#### *3.2.1. Summary of predictor indices*

Due to strong collinearity among predictor indices, bottom depth (*BtmD*) was not included as a continuous predictor in candidate models for station- or transect-based analyses. A binary bottom depth factor, *fBT*, was created using the 100-m depth contour to represent the edge of banks to categorize sample location as occurring over banks (<100 m) or troughs (≥100 m). The depth factor accounted for observed differences in density relative to bottom depth (Fig. 2) by allowing the intercept value to be lower for trough samples.

Station-based oceanographic samples collected in summer and fall 2011 and summer 2013 are summarized (Fig. 4) relative to capelin presence (box plots) and density (scatter plots). Visual inspection of scatter plots indicated that there were categorical differences in oceanographic samples relative to capelin non-zero densities between banks (*n*=18) and troughs (*n*=9). Capelin non-zero density samples occurred in waters over banks that had relatively cooler surface (*TmpS*) and warmer bottom (*TmpB*) temperatures, reduced thermal stratification (*TmpD*), lower bottom salinity (*SalB*), and higher integrated chlorophyll-*a* concentrations (*Chla*) compared to oceanographic conditions in troughs. Oceanographic conditions in summers were similar between years. In fall 2011, water temperatures were cooler and more stratified over banks and troughs compared to summer of that year.

Characterization of oceanographic conditions where capelin were present relative to conditions at stations where they were absent was limited to observations from summer 2013 due to the high occurrence of unknown backscatter in both 2011 surveys. Visual inspection of boxplots (Fig. 4) indicated that stations where capelin were present had relatively warmer bottom temperatures, lower thermal stratification, and higher chlorophyll concentrations compared to stations where capelin were absent.

Distributions of capelin density in summer 2013 and acoustic-based, biological predictors at transect- and station-based analysis resolutions are shown in Fig. 5, along with summary boxplots of continuous biological predictors and the *Edge* index to highlight differences in predictor values based on the presence or absence of capelin, analysis resolution, and bottom depth strata. When all samples for each predictor index were pooled across bottom depths, relationships between median values for *Edge*, *Pred*, and *Prey* indices and capelin presence/absence are similar at both analysis resolutions: median values for *Edge* were higher where capelin were present compared to where they were absent, and median values for *Pred* and *Prey* were lower where capelin were present. In contrast, the relationship between median values for *Comp* based on capelin presence/absence switched between analysis resolutions: median *Comp* was higher where capelin were present at the transect resolution, but was

lower at the station resolution. Overlap in the 25$^{th}$ and 75$^{th}$ quantiles for each predictor index among presence and absence values highlight the limitations of comparing samples pooled across all bottom depths. When samples were partitioned between bottom depth categories, differences in median values relative to capelin presence/absence were more pronounced at both resolutions for *Edge* samples in troughs and *Comp* samples over banks. The relationship between medians for *Pred* relative to capelin presence/absence was opposite between banks (higher where capelin were present) and troughs (lower where capelin were present). Differences in medians for *Prey* based on capelin presence/absence were lower when samples were stratified by depth, although the difference in median values between banks/trough samples was larger and the range between 25$^{th}$ and 75$^{th}$ quantile values was reduced. These results indicate that relationships between capelin distributions, biological predictors, and the *Edge* index are sensitive to analysis resolution and/or bottom depth strata.

*3.2.2. Station-based analyses*

For all station-based candidate models, derived quantities for the geostatistical range and marginal standard deviation for the spatial random field were near zero. Spatial random effects were subsequently dropped and a generalized linear model (GLM) was used to estimate the probability of capelin occurrence and density. Visual inspection of residuals plotted against latitude and longitude indicated no apparent spatial structure (not shown).

Station-based analyses of capelin occurrence among 5 candidate models (Table 3) indicate that model s.P.1 is the best model based on AICc selection. Model s.P.1 predicts that the probability of capelin occurrence in summer 2013 increases with warmer bottom temperature (*TmpB*) and greater distance from the edge of banks (*Edge*), explaining 17.6% of observed variance. In all other candidate models bottom depth factor (*fBT*) is retained in the final model indicating that occurrence probabilities are consistently lower in troughs. The *Edge* predictor is retained in three of four candidate models that

included *fBT*, including two models where predator density (*Pred*) is also retained in the final model. Thermal stratification (*TmpD*) and integrated chlorophyll-*a* concentration (*Chla*) are retained in their final models. When included, higher occurrence probabilities are associated with increasing *Edge*, *TmpB*, *Chla*, and *Pred* and decreasing *TmpD* values. Potential predictors that are not retained in the final model for any candidate models include surface temperature (*TmpS*), bottom salinity (*SalB*), competitor density (*Comp*), and prey density (*Prey*).

The relative importance of each standardized (i.e. subtract mean and divide by standard deviation) predictor within a model (i.e. explanatory power) was examined by comparing the magnitude of its coefficient with estimates of other predictors included in the final model. In model s.P.1, *TmpB* explains more of the variance relative to *Edge* ($\hat{\beta}_{TmpB}$=0.95; $\hat{\beta}_{Edge}$=0.56), whereas in models s.P.2-4, *Edge* has a relatively higher explanatory power compared to the other continuous predictors. To evaluate model fits, predicted occurrence probabilities were visually compared to observed probabilities (Fig. 6A). Model s.P.1 tended to over-estimate occurrence probabilities where capelin were least likely to occur (minimum predicted encounter probability = 0.09), but otherwise model fit was acceptable given that the diagnostic is sensitive to small sample sizes at the resolution (0.1 increments of occurrence probabilities) used in the analysis.

In contrast to occurrence model results, models s.D.1 and s.D.2 (hereafter s.D.1/2 when referring to the same final model for both candidate models) predict that capelin density will be higher at stations located closer to the edge of banks; explaining 43.1% of observed variance. Model s.D.6 is the next best model based on AICc, explaining 37.3% of variance and predicts an increase in density when thermal stratification (*TmpD*) is reduced. Comparison of model diagnostics indicates that model s.D.1/2 fits the data relatively well compared to model s.D.6, but consistently overestimated density (Fig. 6B-D). Model diagnostics reveal that predicted densities from other candidate models fit the observed data poorly (not shown). This includes models s.D.3-4 (*TmpS* and *fBT* in final model), which

488 explains 41.4% of the variation and has a similar CV value as the best model ($CV_{s.D.1-2}$=1.09; $CV_{s.D.3-4}$=1.1),
489 but did not fit the data well as revealed in the diagnostics for these models.
490
491 *3.2.3. Transect-based analyses*
492 Model selection for transect-based analysis of capelin occurrence (Table 4) indicates that the
493 best model to estimate probability of occurrence includes spatial random effects, a negative bottom
494 depth factor (i.e. occurrence probabilities lower for trough samples), and positive coefficients for
495 predator density (*Pred*), competitor density (*Comp*), and prey density (*Prey*). Model t.P.1 explains 71.8%
496 of variance in occurrence, with nearly all of the explained variance attributed to spatial effects. The
497 minor contribution of biological predictors is evident by low parameter estimate values ($\hat{\beta}$ range 0.2 to
498 0.26), indicating no clear differences in explanatory power. Visual inspection of the diagnostic plot (Fig.
499 6A) indicates that predicted occurrence probabilities were over-estimated where capelin are least likely
500 to occur, but overall the model fit to the observed distribution is acceptable based on all observed
501 occurrence probabilities falling within the 95% confidence interval for predicted occurrence
502 probabilities.
503 At the transect-based analysis resolution, the best model based on AICc of capelin density (t.D.1)
504 includes spatial random effects, a negative bottom depth factor, and the edge of banks (*Edge*), *Pred*, and
505 *Comp* predictors (Table 4). Model t.D.1 explains 79.6% of observed variance and predicts that densities
506 will increase with closer proximity to bank edges and increasing densities of predators and competitors.
507 Unlike the transect-based occurrence model, there are clear differences in explanatory power among
508 predictor indices. The *Edge* predictor ($\hat{\beta}$=-0.61) contributes more to model predictions of capelin
509 densities than biological predictors *Pred* ($\hat{\beta}$=0.1) and *Comp* ($\hat{\beta}$=0.18). Visual inspection of diagnostic
510 plots (Fig. 6B-D) shows a relatively good fit and no trends in the residuals, but that predicted densities
511 are under-estimated.

512 Given the observed differences in capelin densities over banks compared to those in troughs
513 (Fig. 4-5), the spatial GLMM used for transect-based analyses was modified by expanding Eq. 2 to
514 estimate an interaction between the *Edge* and biological predictors with the bottom depth factor
515 (hereafter depth-stratified model). By allowing intercepts and slopes to vary based on the sample's
516 location, the model accounts for depth-specific differences in the influence of each covariate. Model
517 selection results for the depth-stratified model are mixed (Table 5). Explained deviance and model fit
518 improved in the depth-stratified occurrence model (t.P_bt), but increases are modest compared to the
519 simpler transect-based model t.P.1 (explained deviance increased by 0.1%; AICc value increased by 6.2).
520 In comparison, for the depth-stratified density model the explained deviance increased by 0.4%, the
521 residual CV was reduced by 0.01, and the AICc value was reduced by 13.7. The improvement in AICc is
522 notable considering the metric penalizes the addition of four new parameters in the depth-stratified
523 model (Hardin and Hilbe, 2007).

524 The depth-stratified models also highlight differences in the relative influence of predictors on
525 capelin distributions between bottom depth strata. Parameter estimates for the competitor density
526 (*Comp*) predictor over banks are larger and have smaller standard errors in both occurrence ($Comp_b$:
527 $\hat{\beta}$=0.27, SE=0.18; $Comp_t$: $\hat{\beta}$=0.07, SE=0.22) and density ($Comp_b$: $\hat{\beta}$=0.26, SE=0.13; $Comp_t$: $\hat{\beta}$=-0.06,
528 SE=0.14) models, indicating that the explanatory power of *Comp* is both stronger and more stable over
529 banks compared to troughs. The predator density (*Pred*) predictor over banks is retained in the
530 occurrence model but dropped from the density model, while the *Pred* predictor for troughs is dropped
531 from the occurrence model but retained in the density model. The explanatory power of the depth-
532 stratified *Pred* predictor is greater relative to the other covariates in the depth-stratified model
533 compared to the simpler, transect-based models (Tables 4-5).

534

535 *3.3. Resolution-dependent model differences*

Comparison of model selection results from station- to transect-based analyses for candidate models with continuous predictors revealed differences in retained predictors and the relative magnitudes of their estimates (Fig. 7). It should be noted that the best station-based occurrence model, s.P.1, was not included in this comparison because it included oceanographic predictors. The final station-based occurrence model (s.P.3) retains *Edge* and *Pred*, with the *Edge* variable explaining more of the variance ($\hat{\beta}_{Edge}$=0.78, $\hat{\beta}_{Pred}$=0.31). In contrast, the final transect-based occurrence model (t.P.1) does not retain the *Edge* predictor and there is no difference in the explanatory power among biological predictors *Pred*, *Comp*, and *Prey*.

In density models s.D.2 and t.D.1, *Edge* is retained in final models at both resolutions. Although *Pred* and *Comp* are also retained in the transect-based density model, the explanatory power of the *Edge* predictor is higher ($\hat{\beta}_{Edge}$=-0.61, $\hat{\beta}_{Pred}$=0.1, $\hat{\beta}_{Comp}$=0.18). Overall, these results indicate that *Edge* is important at both resolutions, while the influence of biological predictors is sensitive to analysis resolution.

## 4. DISCUSSION

### *4.1. Relative importance of environmental factors*

Model selection results indicate that the probability of capelin occurrence increases in waters with warmer bottom temperatures and higher chlorophyll concentrations, and that capelin density increases at shorter distances from bank edges and reduced thermal stratification. These results demonstrate that capelin distributions are influenced by processes associated with increased vertical mixing in association with bathymetry. Warmer bottom temperatures, increased chlorophyll concentrations, and reduced thermal stratification are all characteristic of increased vertical mixing in

the northern GOA (Stabeno et al., 2016). Throughout summer, primary production is sustained over the CGOA shelf by strong tidal pumping that supplies nutrients from bottom water originating in troughs to the euphotic zone over banks and near their edges (Cheng et al., 2012; Ladd et al., 2005; Mordy et al., 2016). Mixing and primary production are locally enhanced over and near CGOA banks due to increased current velocities that result from the interaction of tidal currents and the Alaska Coastal Current (ACC) with steep walls along troughs (Mordy et al., 2016; Stabeno et al., 2016). Primary production is further enhanced by increases in ACC transport, which supplies shelf waters with iron originating primarily from river discharge (Cheng et al., 2012; Lippiatt et al., 2010; Stabeno et al., 2004).

The distribution and species composition of zooplankton that capelin consume over the GOA shelf is correlated with water temperature, salinity, and primary production (Coyle et al., 2013; Coyle and Pinchuk, 2005, 2003). Capelin diets primarily consist of copepods and euphausiids (Logerwell et al., 2010; Wilson et al., 2006), whose distributions are associated with different oceanographic and bathymetric features. Neritic copepods (e.g. *Pseudocalanus* spp., *Metridia pacifica*, and *Calanus marshallae*) are associated with ACC waters over the inner shelf while oceanic copepods (e.g. *Neocalanus cristatus* and *Eucalanus bungii*) primarily occur over the outer shelf and deeper slope waters (Coyle and Pinchuk, 2005). Euphausiids are also associated with the ACC (i.e. *Thysanoessa* spp.) and oceanic waters over the outer shelf and slope (*Euphausia pacifica*), but primarily in waters greater than 100 m depth (Pinchuk et al., 2008). Acoustics-based distributions of euphausiids across the GOA shelf indicate that they concentrate over deeper waters in troughs and sea valleys (Simonsen et al., 2016), characterized by higher current velocities compared to shallow banks and coastal waters (Wilson et al., 2009).

Our findings build on earlier studies over the CGOA shelf that attribute variability in capelin distributions to a variety of factors. Hollowed et al. (2007) hypothesized that capelin distributions in Barnabas and Chiniak Troughs were limited by oceanographic conditions. Logerwell et al. (2007)

584 suggested that capelin preferred warmer, well-mixed waters associated with the ACC that contained
585 large copepod species in late-summer when age-1 and -2 pollock were not present to compete for
586 shared prey resources. This study's station-based analyses support Logerwell et al.'s (2007) hypothesis
587 that capelin were more likely to occur in well-mixed and productive shallow waters, but we believe that
588 this association is characteristic for distributions of age-1 capelin. Our results also indicate that age-2+
589 capelin primarily occupied cold, stratified waters in troughs, similar to distributions observed by
590 Hollowed et al. (2007), and Logerwell et al. (2010, 2007).

591 Results of this study do not support the hypothesis of Logerwell et al. (2010) that capelin
592 distributions are influenced by competitive interactions with age-0 pollock. Field observations
593 combined with diet analyses for capelin and age-0 pollock led Logerwell et al. (2010) to speculate that all
594 but the largest capelin were being outcompeted by age-0 pollock for euphausiids, and that capelin
595 switched to alternate prey and expanded their distributions to match the wide availability of copepods.
596 In contrast, this study's transect-based model results showed a positive relationship between capelin
597 and age-0 pollock, and that the competitor index was not retained in any station-based models. This
598 result indicates that capelin distributions were not negatively influenced by competitive interactions
599 between species at either analysis resolution, and that the importance of the relationship differed
600 between analysis resolutions. Interestingly, in both the depth-stratified, transect-based occurrence and
601 density models, the explanatory power of the competitor index was higher over banks, compared to
602 troughs. The positive relationship between capelin and age-0 pollock suggests that these species were
603 likely responding to similar environmental cues over banks, which may be associated with utilization of
604 shared resources that are not limiting. Age-0 pollock primarily consumed small and large copepods over
605 the CGOA shelf in August 2013 (Moss et al., 2016a), consistent with the observation that diets of capelin
606 less than 9.5 cm standard length were also dominated by copepods in 2004 and 2005 (Logerwell et al.,
607 2010). A lack of difference between capelin vertical distribution patterns in 2011 when pollock were

608 absent and during 2013 when pollock were abundant (McGowan et al., 2016) further suggests that
609 capelin vertical movements were not influenced by competitive interactions with pollock in summer.
610 Our results do not indicate how the relationship might change in the fall when euphausiids would
611 comprise a greater proportion of diets for larger age-0 pollock as reported in Logerwell et al. (2010) and
612 Wilson et al. (2006).

613 Comparison of bank and trough parameter estimates for the competitor index in the depth-
614 stratified transect-based models highlight differences among capelin age-classes. The depth-stratified
615 model was able to differentiate a strong, positive relationship between age-0 pollock and age-1 capelin
616 density over banks, and a relatively weak, negative relationship with larger, age-2+ capelin in troughs.
617 Age-specific differences in prey and predator predictors also support hypothesized differences in prey
618 preferences. The positive relationship between capelin density and macrozooplankton was relatively
619 weak over banks compared to troughs. This is consistent with age-2+ capelin consuming larger
620 zooplankton prey in troughs, such as euphausiids (Logerwell et al., 2010; Wilson et al., 2006). Capelin
621 were positively related to predator density in transect-based models for occurrence and density, with
622 predator density more influential in predicting capelin densities in troughs. The positive relationship
623 between the predator index and capelin in both station- and transect-based models indicates that
624 capelin distributions do not appear to be influenced by predation or avoidance behavior.

625 Observed differences in capelin length frequency distributions and literature-based, diet
626 information leads to the hypothesis that in summer age-1 capelin concentrate over shallow banks on
627 the CGOA shelf in waters that are well-mixed and more productive to feed on abundant copepods. Age-
628 2+ capelin primarily occupy deeper trough waters where they consume larger prey, such as euphausiids.
629 Distributions of both age groups over the shelf will likely be driven by changes in the composition and
630 availability of their prey. Our results are consistent with “bottom-up control” (e.g. Speckman et al.,
631 2005), but do not indicate specific mechanisms that determine capelin distributions. To better

understand mechanisms responsible for structuring capelin distributions over the GOA shelf, future studies should include age-based differences in distribution patterns (e.g. following Thorson et al., *In press*, for pollock).

Observations from inshore Alaskan waters and Atlantic capelin populations support the hypothesis that capelin respond to oceanographically-induced changes in prey availability. During summer in southeast Alaska, Arimitsu et al. (2008) reported high densities of immature capelin and spawning adults in productive areas of Glacier Bay that were adjacent to edges of deep basins where interactions between tidal currents and shallow sills generate upwelling. In the North Atlantic, age-1 and -2 capelin were distributed in well-mixed, zooplankton-rich waters over the shelf north of Iceland and on the East Greenland plateau, where primary production is sustained by infusions of warm, nutrient-rich water from the Atlantic (Vilhjálmsson, 2002). In the Barents Sea, immature capelin distributions are associated with the receding edge of sea ice throughout summer, where primary and secondary production occur in waters near the ice edge as it retreats northward (Gjøsæter, 1998).

*4.2. Analysis resolution and modeling approach*

Integrating continuous acoustic sampling across bathymetric gradients with systematic oceanographic sampling at discrete stations allowed us to identify scale-dependent relationships among environmental factors and capelin distributions. Previous studies have identified scale-dependent relationships between capelin and their predators, including cod (Rose and Leggett, 1990) and seabirds (Fauchald et al., 2000; Piatt, 1990). Recognizing that capelin abundance was relatively high in 2013 compared to 2011 (McGowan et al., 2016) and that the magnitude and direction of predator-prey relationships have the potential to be both density- and scale-dependent (Fauchald, 2009), future studies should examine how interannual variability in capelin abundance is related to distributions of their prey, predators, and potential competitors.

Examination of autocorrelated samples at a high (0.5 km) resolution was possible by treating spatial covariance among autocorrelated samples as a random effect in the spatial GLMM. Traditional approaches for analyzing spatial data typically resample response variable data at coarser resolutions, or use nested hierarchical approaches to remove spatial structure to meet assumptions of independence among samples (Fauchald et al., 2000; Legendre and Fortin, 1989; Zuur et al., 2009). Treating spatial covariance as a random effect allowed unobserved extrinsic (e.g. density-independent response to environmental conditions) and/or intrinsic (e.g. density-dependent schooling behavior) processes that caused autocorrelation among samples to be included in the prediction of capelin distributions (Beale et al., 2010; Thorson et al., 2015a, 2015b). The spatial GLMM facilitated examination of the relative importance of environmental factors to explain variability in capelin distributions at biologically relevant spatial scales. Compared to station-based models, the spatial GLMM coefficient values for fixed covariates were smaller in transect-based models as spatial residuals accounted for most of the variance. Smaller parameter values was an expected result as variance related to spatial relationships are partitioned between random and fixed effects in the spatial GLMM, while variance is only attributed to fixed covariates in the non-spatial GLM used in conventional station-based analyses (Beale et al., 2010).

Our findings also highlight the importance of considering length/age-based differences in distributions that may be influenced by different physical and biological processes. Given differences in length frequency distributions from trawl samples between banks and troughs, we inferred that age-1 fish primarily occupied banks while age-2+ fish were more prevalent in troughs. Age-based differences in capelin distributions and apparent differences in oceanographic properties between banks and troughs suggest that capelin occupy different water masses based on life stage. It is not surprising that there were differences in the magnitudes and signs of parameter estimates for bank and trough coefficients in the depth-stratified transect-based models. Differences among the bank/trough

coefficients indicate that the relative influence of environmental factors on the distribution of capelin is associated with the age-structure of the observed population. Stratifying predictor indices by bottom depth categories also reduced collinearity among oceanographic (not shown) and biological predictors, facilitating model selection of depth-stratified candidate models that were comprised of all potential predictors. This streamlined identification of covariates that best fit the data and explained the most deviance in capelin distributions by eliminating the step of comparing multiple candidate models.

*4.3. Study limitations*

Examining the relationship between physical and biological factors on capelin distributions was limited in this study to a single year. Oceanographic conditions over the GOA shelf in 2013 are believed to represent average conditions (Hopcroft et al., 2016), while capelin densities were above average relative to other time series (Ormseth, 2012). It is premature to speculate how capelin would respond to environmental perturbations at different population abundance levels. Nonetheless, our results provide guidance to future studies that investigate how capelin distributions are influenced by environmental variability as a function of data resolution and how model structure potentially impacts analysis results.

Results of this study may have been influenced by the sampling and environmental indices used as potential predictors. The relatively high proportion of acoustic data classified as unknown in summer and fall 2011 increased uncertainty when differentiating absences (i.e. true zeroes) from undetected capelin (i.e. sampling zeroes). The analysis of 2011 survey data was limited to only samples where capelin were present when quantifying changes in density relative to bottom depth by year and season. Exclusion of 2011 data from regression analyses prevented running depth-stratified models at the station resolution due to insufficient sample sizes.

703 Differences in capelin density between banks and troughs may also be due to fish that were
704 located deeper in the water column with corresponding lower acoustic target strengths (TS). The
705 vertical position of capelin was significantly shallower over banks (median center of mass = 55 m)
706 compared to troughs (median center of mass ≥ 110 m) in both years (McGowan et al., 2016). Because
707 capelin have a physostomous swimbladder (Fahlén, 1968), compression of their swimbladder at greater
708 depths results in gas loss than cannot be replaced at depth, thereby reducing their backscattering cross-
709 section ($\sigma_{bs}$, e.g. Gorska and Ona, 2003), and resulting in lower acoustic densities (Blaxter and Batty,
710 1990; Jørgensen, 2003). A depth-dependent TS correction factor is not available for capelin (De Robertis
711 et al., 2017b), and we did not adjust acoustic densities for the potential depth bias. Nonetheless, we
712 believe that the magnitude of the bias is relatively low and constant due to differences in acoustic
713 observations of school morphology and trawl catches. Pronounced morphological differences were
714 observed in capelin densities: schools located over banks were 5 to 20 m in height compared to small,
715 discrete schools in troughs that were rarely greater than 5 m in height. Trawl samples support this
716 conclusion, as midwater catches were consistently higher over banks (1000s to 10,000s of capelin per
717 haul) than in troughs (100s to 1000s of fish), even though selectivity of the Cantrawl net is higher for
718 capelin larger than 80 mm (De Robertis et al., 2017a).
719 The acoustic-based macrozooplankton index was intended as a proxy for relative densities of
720 zooplankton biomass, but was likely biased toward species primarily consumed by age-2+ capelin.
721 Copepods are important prey for age-1 and -2+ capelin (Logerwell et al., 2010; Wilson et al., 2006), but
722 are relatively weak scatterers at the frequencies (38 and 120 kHz) used in this study (Matsukura et al.,
723 2009). Larger zooplankton species that are consumed by age-2+ capelin, such as euphausiids and
724 amphipods, have higher TS values (De Robertis et al., 2010; Kang et al., 2002; Murase et al., 2009). As a
725 result, measures of acoustic copepod density may be masked by backscatter associated with
726 euphausiids and amphipods, making the macrozooplankton index a poor measure of copepod biomass

(Witteveen et al., 2015). Net-based measures of copepod biomass (*sensu* Coyle and Pinchuk, 2003; Wilson, 2009) were not available at approximately half of sampled stations and were not included in this analysis.

Predator density was based on the piscivores classification category from McGowan et al. (2016), where backscatter originated from large aggregations of adult groundfish, but also included age-1+ groundfish that may forage at a similar trophic level as capelin. Trawl samples indicated that age-2+ capelin often co-occurred with juvenile pollock (likely age-1 and -2 based on lengths, Brodeur and Wilson, 1996) in troughs. Even though age-1 and -2 pollock may consume similar prey resources as age-2+ capelin (Logerwell et al., 2007), they were classified as piscivorous fish due to insufficient trawl samples to differentiate them from larger pollock. Positive covariance between capelin distributions and the predator index may be partially attributed to both species responding to similar foraging cues, not unlike the positive relationship observed between age-1 capelin and age-0 pollock over banks. This potential bias does not affect our findings that capelin distributions were not likely influenced by predator density in summer 2013.

### *4.4. Summary*

This study's integrated sampling design allowed us to quantify the relative importance of physical and biological environmental factors when characterizing distributions of a mobile, pelagic species at two spatial resolutions. Capelin distributions on the CGOA shelf were primarily concentrated over shallow submarine banks in areas of enhanced vertical mixing and primary production. In summer 2013, capelin were predicted to occur across the shelf in waters with relatively warmer bottom temperatures, with highest densities occurring over banks or in close proximity to bank edges. The strong association between capelin density and proximity to banks was apparent at both analysis resolutions. The relative importance of biological indices as predictors of occurrence and density

751 differed between sample resolutions and bottom depth strata. Our findings indicate that in summer
752 age-1 capelin concentrate over shallow banks on the CGOA shelf in waters that are well-mixed and more
753 productive (Cheng et al., 2012; Mordy et al., 2016), and where copepods are likely to be abundant
754 (Coyle et al., 2013). Age-2+ capelin primarily occupy deeper trough waters where they likely consume
755 larger zooplankton prey, such as euphausiids (Wilson et al., 2009). Recognizing that the influence of
756 environmental processes on capelin distributions may vary based on the age-structure of the observed
757 population, we recommend that studies investigating mechanisms responsible for structuring capelin
758 distributions include size- and/or age-specific analyses.
759

760 **Acknowledgements**

761 We thank Captain Ray Haddon, the crew of the F/V *Northwest Explorer*, and members of the
762 scientific parties on GOAIERP cruises in 2011 and 2013. Special thanks to D. Barbee and J. Nomura
763 (University of Washington, UW) for assisting with the collection and processing of acoustic data during
764 the 2011 surveys. We greatly appreciate our GOAIERP partners providing processed oceanographic
765 data, including S. Danielson (University of Alaska-Fairbanks), M. Sullivan (NOAA Pacific Marine
766 Environmental Lab, PMEL), and S. Strom (University Western Washington), and wish to thank J. Gann
767 and K. Cieciel (NOAA Alaska Fisheries Science Center, AFSC, Auke Bay Laboratories) for assistance with
768 CTD data. Members of the University of Washington Fisheries Acoustic Lab are thanked for assistance
769 with software and countless conversations that improved this study's analysis, with special thanks to E.
770 Phillips for assistance with GLM and GLMM analyses. The analysis design was greatly improved from
771 conversations with C. Ladd (NOAA PMEL), P. MacCready (UW), T. Walsworth (UW), and our GOAIERP
772 partners. We thank L. Rogers (NOAA AFSC), L. Ciannelli (Oregon State University) and an anonymous
773 reviewer for their thoughtful comments and suggested revisions to earlier versions of this manuscript.
774 Funding was provided by the North Pacific Research Board (NPRB) and NOAA Fisheries. The findings and

conclusions in this paper are those of the authors and do not necessarily represent the views of the National Marine Fisheries Service, NOAA. Reference to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA. This paper represents GOAIERP publication #24 and NPRB publication #657.

## 5. REFERENCES

Anderson, P.J., Piatt, J.F., 1999. Community reorganization in the Gulf of Alaska following ocean climate regime shift. Mar. Ecol. Prog. Ser. 189, 117–123.

Andrews, A.G., Strasburger, W.W., Farley, E.V., Murphy, J.M., Coyle, K.O., 2015. Effects of warm and cold climate conditions on capelin (*Mallotus villosus*) and Pacific herring (*Clupea pallasii*) in the eastern Bering Sea. Deep-Sea Res. Part II. https://doi.org/10.1016/j.dsr2.2015.10.008

Arimitsu, M.L., Piatt, J.F., Litzow, M.A., Abookire, A.A., Romano, M.D., Robards, M.D., 2008. Distribution and spawning dynamics of capelin (*Mallotus villosus*) in Glacier Bay, Alaska: A cold water refugium. Fish. Oceanogr. 17, 137–146. https://doi.org/10.1111/j.1365-2419.2008.00470.x

Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J., Elston, D.A., 2010. Regression analysis of spatial data. Ecol. Lett. 13, 246–264. https://doi.org/10.1111/j.1461-0248.2009.01422.x

Blaxter, J.H.S., Batty, R.S., 1990. Swimbladder "behaviour" and target strength. Rapp. Procès-Verbaux La Réun. Cons. Int. Pour Explor. Mer 189, 233–244.

Bradshaw, G.A., Spies, T.A., 1992. Characterizing canopy gap structure in forests using wavelet analysis. J. Ecol. 80, 205–215. https://doi.org/10.2307/2261007

Brodeur, R.D., Wilson, M.T., 1996. A review of the distribution, ecology and population dynamics of age-0 walleye pollock in the Gulf of Alaska. Fish. Oceanogr. 5, 148–166. https://doi.org/10.1111/j.1365-2419.1996.tb00089.x

Brown, E.D., 2002. Life history, distribution, and size structure of Pacific capelin in Prince William Sound and the northern Gulf of Alaska. ICES J. Mar. Sci. 59, 983–996. https://doi.org/10.1006/jmsc.2002.1281

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd ed. Springer, New York, New York, USA.

Carscadden, J.E., Gjøsæter, H., Vilhjálmsson, H., 2013a. Recruitment in the Barents Sea, Icelandic, and eastern Newfoundland/Labrador capelin (*Mallotus villosus*) stocks. Prog. Oceanogr., Norway-Canada Comparison of Marine Ecosystems (NORCAN) 114, 84–96. https://doi.org/10.1016/j.pocean.2013.05.006

Carscadden, J.E., Gjøsæter, H., Vilhjálmsson, H., 2013b. A comparison of recent changes in distribution of capelin (*Mallotus villosus*) in the Barents Sea, around Iceland and in the Northwest Atlantic. Prog. Oceanogr., Norway-Canada Comparison of Marine Ecosystems (NORCAN) 114, 64–83. https://doi.org/10.1016/j.pocean.2013.05.005

Cheng, W., Hermann, A.J., Coyle, K.O., Dobbins, E.L., Kachel, N.B., Stabeno, P.J., 2012. Macro- and micro-nutrient flux to a highly productive submarine bank in the Gulf of Alaska: A model-based analysis of daily and interannual variability. Prog. Oceanogr. 101, 63–77. https://doi.org/10.1016/j.pocean.2012.01.001

Coyle, K.O., Gibson, G.A., Hedstrom, K., Hermann, A.J., Hopcroft, R.R., 2013. Zooplankton biomass, advection and production on the northern Gulf of Alaska shelf from simulations and field observations. J. Mar. Syst. 128, 185–207. https://doi.org/10.1016/j.jmarsys.2013.04.018

Coyle, K.O., Pinchuk, A.I., 2005. Seasonal cross-shelf distribution of major zooplankton taxa on the northern Gulf of Alaska shelf relative to water mass properties, species depth preferences and vertical migration behavior. Deep-Sea Res. Part II. 52, 217–245. https://doi.org/10.1016/j.dsr2.2004.09.025

Coyle, K.O., Pinchuk, A.I., 2003. Annual cycle of zooplankton abundance, biomass and production on the northern Gulf of Alaska shelf, October 1997 through October 2000. Fish. Oceanogr. 12, 327–338. https://doi.org/10.1046/j.1365-2419.2003.00256.x

De Robertis, A., McKelvey, D.R., Ressler, P.H., 2010. Development and application of an empirical multifrequency method for backscatter classification. Can. J. Fish. Aquat. Sci. 67, 1459–1474. https://doi.org/10.1139/F10-075

De Robertis, A., Taylor, K., Williams, K., Wilson, C.D., 2017a. Species and size selectivity of two midwater trawls used in an acoustic survey of the Alaska Arctic. Deep-Sea Res. Part II. 135, 40–50. https://doi.org/10.1016/j.dsr2.2015.11.014

De Robertis, A., Taylor, K., Wilson, C.D., Farley, E.V., 2017b. Abundance and distribution of Arctic cod (*Boreogadus saida*) and other pelagic fishes over the U.S. Continental Shelf of the Northern Bering and Chukchi Seas. Deep-Sea Res. Part II. 135, 51–65. https://doi.org/10.1016/j.dsr2.2016.03.002

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M., Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30, 609–628. https://doi.org/10.1111/j.2007.0906-7590.05171.x

Fahlén, G., 1968. The gas bladder as a hydrostatic organ in *Thymallus thymallus* L., *Osmerus eperlanus* L. and *Mallotus villosus* Müll. Fisk. Skr. Ser. Havunders. 14, 199–228.

Farley, E.V., Murphy, J.M., Wing, B.W., Moss, J.H., Middleton, A., 2005. Distribution, migration pathways, and size of western Alaska juvenile salmon along the eastern Bering Sea shelf. Alsk. Fish. Res. Bull. 11, 15–26.

Fauchald, P., 2009. Spatial interaction between seabirds and prey: review and synthesis. Mar. Ecol. Prog. Ser. 391, 139–151. https://doi.org/10.3354/meps07818

Fauchald, P., Erikstad, K.E., 2002. Scale-dependent predator-prey interactions: the aggregative response of seabirds to prey under variable prey abundance and patchiness. Mar. Ecol. Prog. Ser. 231, 279–291. https://doi.org/10.3354/meps231279

Fauchald, P., Erikstad, K.E., Skarsfjord, H., 2000. Scale-dependent predator-prey interactions: The hierarchical spatial distribution of seabirds and prey. Ecology 81, 773–783. https://doi.org/10.2307/177376

Foote, K.G., Knudsen, H.P., Vestnes, G., MacLennan, D.N., Simmonds, E.J., 1987. Calibration of acoustic instruments for fish density estimation: a practical guide (Cooperative Research Report No. 144). International Council for the Exploration of the Sea.

Francis, R.C., Hare, S.R., Hollowed, A.B., Wooster, W.S., 1998. Effects of interdecadal climate variability on the oceanic ecosystems of the NE Pacific. Fish. Oceanogr. 7, 1–21. https://doi.org/10.1046/j.1365-2419.1998.00052.x

Frank, K.T., Carscadden, J.E., Simon, J.E., 1996. Recent excursions of capelin (*Mallotus villosus*) to the Scotian Shelf and Flemish Cap during anomalous hydrographic conditions. Can. J. Fish. Aquat. Sci. 53, 1473–1486. https://doi.org/10.1139/f96-078

Gjøsæter, H., 1998. The population biology and exploitation of capelin (*Mallotus villosus*) in the Barents Sea. Sarsia 83, 453–496.

Gjøsæter, H., Bogstad, B., Tjelmeland, S., 2009. Ecosystem effects of the three capelin stock collapses in the Barents Sea. Mar. Biol. Res. 5, 40–53. https://doi.org/10.1080/17451000802454866

Gorska, N., Ona, E., 2003. Modelling the acoustic effect of swimbladder compression in herring. ICES J. Mar. Sci. 60, 548–554. https://doi.org/10.1016/S1054-3139(03)00050-X

Hardin, J.W., Hilbe, J.M., 2007. Generalized Linear Models and Extensions, Second Edition. Stata Press, College Station, Texas.

Hjermann, D.Ø., Bogstad, B., Dingsør, G.E., Gjøsæter, H., Ottersen, G., Eikeset, A.M., Stenseth, N.C., 2010. Trophic interactions affecting a key ecosystem component: a multistage analysis of the recruitment of the Barents Sea capelin (*Mallotus villosus*). Can. J. Fish. Aquat. Sci. 67, 1363–1375. https://doi.org/10.1139/F10-064

Hollowed, A.B., Barbeaux, S.J., Cokelet, E.D., Farley, E., Kotwicki, S., Ressler, P.H., Spital, C., Wilson, C.D., 2012. Effects of climate variations on pelagic ocean habitats and their role in structuring forage fish distributions in the Bering Sea. Deep-Sea Res. Part II. 65–70, 230–250. https://doi.org/10.1016/j.dsr2.2012.02.008

Hollowed, A.B., Wilson, C.D., Stabeno, P.J., Salo, S.A., 2007. Effect of ocean conditions on the cross-shelf distribution of walleye pollock (*Theragra chalcogramma*) and capelin (*Mallotus villosus*). Fish. Oceanogr. 16, 142–154. https://doi.org/10.1111/j.1365-2419.2006.00418.x

Hopcroft, R.R., Aguilar-Islas, A.M., Ladd, C., Matarese, A.C., Mordy, C.W., Rember, R., Stabeno, P.J., Strom, S.L., 2016. The role of cross-shelf and along-shelf transports as controlling mechanisms for nutrients, plankton and larval fish in the coastal Gulf of Alaska (NPRB GOA Project G83 & G85 Lower Trophic Level Final Report). North Pacific Research Board, Anchorage, AK, 169 pp.

Horne, J.K., Schneider, D.C., 1994. Analysis of scale-dependent processes with dimensionless ratios. Oikos 70, 201–211. https://doi.org/10.2307/3545631

Ingvaldsen, R.B., Gjøsæter, H., 2013. Responses in spatial distribution of Barents Sea capelin to changes in stock size, ocean temperature and ice cover. Mar. Biol. Res. 9, 867–877. https://doi.org/10.1080/17451000.2013.775450

Jørgensen, R., 2003. The effects of swimbladder size, condition and gonads on the acoustic target strength of mature capelin. ICES J. Mar. Sci. 60, 1056–1062. https://doi.org/10.1016/S1054-3139(03)00115-2

Kang, M., Furusawa, M., Miyashita, K., 2002. Effective and accurate use of difference in mean volume backscattering strength to identify fish and plankton. ICES J. Mar. Sci. 59, 794–804. https://doi.org/10.1006/jmsc.2002.1229

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., Bell, B., 2016. TMB: Automatic differentiation and Laplace approximation. J. Stat. Softw. 70, 1–21. https://doi.org/doi: 10.18637/jss.v070.i05

Kutner, M., Nachtsheim, C., Neter, J., Li, W., 2004. Applied linear statistical models, 5th ed. McGraw Hill.

Ladd, C., Stabeno, P., Cokelet, E., 2005. A note on cross-shelf exchange in the northern Gulf of Alaska. Deep-Sea Res. Part II. 52, 667–679.

Legendre, P., Fortin, M., 1989. Spatial Pattern and Ecological Analysis. Vegetatio 80, 107–138. https://doi.org/10.1007/BF00048036

Lindgren, F., Rue, H., Lindström, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B Stat. Methodol. 73, 423–498. https://doi.org/10.1111/j.1467-9868.2011.00777.x

Lippiatt, S.M., Lohan, M.C., Bruland, K.W., 2010. The distribution of reactive iron in northern Gulf of Alaska coastal waters. Mar. Chem. 121, 187–199.

Logerwell, E.A., Duffy-Anderson, J., Wilson, M., McKelvey, D., 2010. The influence of pelagic habitat selection and interspecific competition on productivity of juvenile walleye pollock (*Theragra*

911 *chalcogramma*) and capelin (*Mallotus villosus*) in the Gulf of Alaska. Fish. Oceanogr. 19, 262–
912 278. https://doi.org/10.1111/j.1365-2419.2010.00542.x
913 Logerwell, E.A., Stabeno, P.J., Wilson, C.D., Hollowed, A.B., 2007. The effect of oceanographic variability
914 and interspecific competition on juvenile pollock (*Theragra chalcogramma*) and capelin
915 (*Mallotus villosus*) distributions on the Gulf of Alaska shelf. Deep-Sea Res. Part II. 54, 2849–2868.
916 https://doi.org/10.1016/j.dsr2.2007.08.008
917 MacLennan, D.N., Fernandes, P.G., Dalen, J., 2002. A consistent approach to definitions and symbols in
918 fisheries acoustics. ICES J. Mar. Sci. J. Cons. 59, 365–369.
919 https://doi.org/10.1006/jmsc.2001.1158
920 Matsukura, R., Yasuma, H., Murase, H., Yonezaki, S., Funamoto, T., Honda, S., Miyashita, K., 2009.
921 Measurements of density contrast and sound-speed contrast for target strength estimation of
922 Neocalanus copepods (*Neocalanus cristatus* and *Neocalanus plumchrus*) in the North Pacific
923 Ocean. Fish. Sci. 75, 1377. https://doi.org/10.1007/s12562-009-0172-3
924 McGowan, D.W., Horne, J.K., Parker-Stetter, S.L., 2016. Variability in species composition and
925 distribution of forage fish in the Gulf of Alaska. Deep-Sea Res. Part II.
926 https://doi.org/10.1016/j.dsr2.2016.11.019
927 Merrick, R.L., Chumbley, M.K., Byrd, G.V., 1997. Diet diversity of Steller sea lions (*Eumetopias jubatus*)
928 and their population decline in Alaska: a potential relationship. Can. J. Fish. Aquat. Sci. 54, 1342–
929 1348. https://doi.org/10.1139/f97-037
930 Mordy, C.W., Stabeno, P.J., Kachel, N.B., Kachel, D., Ladd, C., Zimmermann, M., Doyle, M.J., 2016.
931 Appendix 1: Importance of canyons to the northern Gulf of Alaska ecosystem, in: The role of
932 cross-shelf and along-shelf transports as controlling mechanisms for nutrients, plankton and
933 larval fish in the coastal Gulf of Alaska. NPRB GOA Project G83 & G85 Lower Trophic Level Final
934 Report, North Pacific Research Board, Anchorage, AK, 45-77.
935 Moss, J.H., Heintz, R.A., Zaleski, M., 2016a. Chapter 10 - Interannual and regional variability in the
936 distribution, feeding, and energetic health of age-0 walleye pollock (*Gadus chalcogrammus*) and
937 pacific cod (*Gadus macrocephalus*) in the Gulf of Alaska, in: Surviving the Gauntlet: a
938 comparative study of the pelagic, demersal, and spatial linkages that determine groundfish
939 recruitment and diversity in the Gulf of Alaska ecosystem. NPRB GOA Project G84 Upper Trophic
940 Level Final Report, North Pacific Research Board, Anchorage, AK, 296-313.
941 Moss, J.H., Shotwell, S.K., Heintz, R.A., Atkinson, S., Debenham, C., Fournier, W., Golden, N., Heifetz, J.,
942 Mueter, F.J., Pirtle, J.L., Reid, J.A., Slater, L., Sreenivasan, A., Will, A., Zaleski, M., Zimmermann,
943 M., 2016b. Surviving the Gauntlet: a comparative study of the pelagic, demersal, and spatial
944 linkages that determine groundfish recruitment and diversity in the Gulf of Alaska ecosystem
945 (NPRB GOA Project G84 Upper Trophic Level Final Report). North Pacific Research Board,
946 Anchorage, AK.
947 Murase, H., Ichihara, M., Yasuma, H., Watanabe, H., Yonezaki, S., Nagashima, H., Kawahara, S.,
948 Miyashita, K., 2009. Acoustic characterization of biological backscatterings in the Kuroshio-
949 Oyashio inter-frontal zone and subarctic waters of the western North Pacific in spring. Fish.
950 Oceanogr. 18, 386–401. https://doi.org/10.1111/j.1365-2419.2009.00519.x
951 Olafsdottir, A.H., Rose, G.A., 2012. Influences of temperature, bathymetry and fronts on spawning
952 migration routes of Icelandic capelin (*Mallotus villosus*). Fish. Oceanogr. 21, 182–198.
953 https://doi.org/10.1111/j.1365-2419.2012.00618.x
954 Ona, E., Mitson, R.B., 1996. Acoustic sampling and signal processing near the seabed: the deadzone
955 revisited. ICES J. Mar. Sci. J. Cons. 53, 677–690. https://doi.org/10.1006/jmsc.1996.0087
956 Ormseth, O., 2012. Appendix 2. Preliminary assessment of forage species in the Gulf of Alaska (Stock
957 Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska).
958 North Pacific Fishery Management Council, 605 W. 4th Avenue, Suite 306, Anchorage, AK 99301.

Parker-Stetter, S., Urmy, S., Horne, J., Eisner, L., Farley, E., 2016. Factors affecting summer distributions of Bering Sea forage fish species: assessing competing hypotheses. Deep Sea Res. Part II Top. Stud. Oceanogr. https://doi.org/10.1016/j.dsr2.2016.06.013i

Piatt, J.F., 1990. The aggregative response of common murres and Atlantic puffins to schools of capelin. Stud. Avian Ecol. 14, 36–51.

Piatt, J.F., Anderson, P.J., 1996. Response of common murres to the "Exxon Valdez" oil spill and long-term changes in the Gulf of Alaska marine ecosystem, in: American Fisheries Society Symposium. Presented at the Proceedings of the Exxon Valdez Oil Spill Symposium: held at Anchorage, Alaska, USA, 2-5 February 1993, American Fisheries Society, Bethesda, MD, pp. 720–737.

Pinchuk, A.I., Coyle, K.O., Hopcroft, R.R., 2008. Climate-related variability in abundance and reproduction of euphausiids in the northern Gulf of Alaska in 1998–2003. Prog. Oceanogr. 77, 203–216. https://doi.org/10.1016/j.pocean.2008.03.012

R Core Development Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rose, G.A., 2005. Capelin (*Mallotus villosus*) distribution and climate: a sea "canary" for marine ecosystem change. ICES J. Mar. Sci. J. Cons. 62, 1524–1530. https://doi.org/10.1016/j.icesjms.2005.05.008

Rose, G.A., Leggett, W.C., 1990. The importance of scale to predator-prey spatial correlations: an example of Atlantic fishes. Ecology 71, 33–43. https://doi.org/10.2307/1940245

Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J. R. Stat. Soc. Ser. B Stat. Methodol. 71, 319–392. https://doi.org/10.1111/j.1467-9868.2008.00700.x

Schneider, D.C., 1994. Quantitative ecology: spatial and temporal scaling. Academic Press, San Diego.

Simmonds, E.J., MacLennan, D.N., 2005. Fisheries Acoustics: Theory and Practice, 2nd ed. Blackwell Science, Oxford, UK.

Simonsen, K.A., Ressler, P.H., Rooper, C.N., Zador, S.G., 2016. Spatio-temporal distribution of euphausiids: an important component to understanding ecosystem processes in the Gulf of Alaska and eastern Bering Sea. ICES J. Mar. Sci. 73, 2020–2036. https://doi.org/10.1093/icesjms/fsv272

Skaug, H.J., Fournier, D.A., 2006. Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. Comput. Stat. Data Anal. 51, 699–709. https://doi.org/10.1016/j.csda.2006.03.005

Speckman, S.G., Piatt, J.F., Minte-Vera, C.V., Parrish, J.K., 2005. Parallel structure among environmental gradients and three trophic levels in a subarctic estuary. Prog. Oceanogr. 66, 25–65. https://doi.org/10.1016/j.pocean.2005.04.001

Springer, A.M., Speckman, S.G., 1997. A forage fish is what? Summary of the symposium (No. AK-SG-97-01), Forage Fishes in Marine Ecosystems. Alaska Sea Grant College Program, Fairbanks, Alaska, USA.

Stabeno, P., Bond, N., Hermann, A., Kachel, N., Mordy, C., Overland, J., 2004. Meteorology and oceanography of the Northern Gulf of Alaska. Cont. Shelf Res. 24, 859–897.

Stabeno, P.J., Bell, S., Cheng, W., Danielson, S., Kachel, N.B., Mordy, C.W., 2016. Long-term observations of Alaska Coastal Current in the northern Gulf of Alaska. Deep-Sea Res. Part II. https://doi.org/10.1016/j.dsr2.2015.12.016

Stabeno, P.J., Bond, N.A., Kachel, N.B., Ladd, C., Mordy, C.W., Strom, S.L., 2015. Southeast Alaskan shelf from southern tip of Baranof Island to Kayak Island: Currents, mixing and chlorophyll-*a*. Deep-Sea Res. Part II. https://doi.org/10.1016/j.dsr2.2015.06.018

Strom, S.L., Fredrickson, K.A., Bright, K.J., 2015. Spring phytoplankton in the eastern coastal Gulf of Alaska: Photosynthesis and production during high and low bloom years. Deep-Sea Res. Part II. https://doi.org/10.1016/j.dsr2.2015.05.003

Thorson, J.T., Ianelli, J.N., Kotwicki, S., 2017. The relative influence of temperature and size-structure on fish distribution shifts: A case-study on Walleye pollock in the Bering Sea. Fish Fish. 18, 1073–1084. https://doi.org/10.1111/faf.12225

Thorson, J.T., Shelton, A.O., Ward, E.J., Skaug, H.J., 2015a. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES J. Mar. Sci. J. Cons. 72, 1297–1310. https://doi.org/10.1093/icesjms/fsu243

Thorson, J.T., Skaug, H.J., Kristensen, K., Shelton, A.O., Ward, E.J., Harms, J.H., Benante, J.A., 2015b. The importance of spatial models for estimating the strength of density dependence. Ecology 96, 1202–1212. https://doi.org/10.1890/14-0739.1

Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bull. Am. Meteorol. Soc. 79, 61–78. https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2

Vilhjálmsson, H., 2002. Capelin (*Mallotus villosus*) in the Iceland–East Greenland–Jan Mayen ecosystem. ICES J. Mar. Sci. J. Cons. 59, 870–883. https://doi.org/10.1006/jmsc.2002.1233

Waite, J.N., Mueter, F.J., 2013. Spatial and temporal variability of chlorophyll-*a* concentrations in the coastal Gulf of Alaska, 1998–2011, using cloud-free reconstructions of SeaWiFS and MODIS-Aqua data. Prog. Oceanogr. 116, 179–192. https://doi.org/10.1016/j.pocean.2013.07.006

Williams, K., Punt, A.E., Wilson, C.D., Horne, J.K., 2011. Length-selective retention of walleye pollock, *Theragra chalcogramma*, by midwater trawls. ICES J. Mar. Sci. J. Cons. 68, 119–129. https://doi.org/10.1093/icesjms/fsq155

Wilson, M.T., 2009. Ecology of small neritic fishes in the western Gulf of Alaska. I. Geographic distribution in relation to prey density and the physical environment. Mar. Ecol. Prog. Ser. 392, 223–237. https://doi.org/10.3354/meps08160

Wilson, M.T., Jump, C.M., Buchheister, A., 2009. Ecology of small neritic fishes in the western Gulf of Alaska. II. Consumption of krill in relation to krill standing stock and the physical environment. Mar. Ecol. Prog. Ser. 392, 239–251. https://doi.org/10.3354/meps08237

Wilson, M.T., Jump, C.M., Duffy-Anderson, J.T., 2006. Comparative analysis of the feeding ecology of two pelagic forage fishes: capelin *Mallotus villosus* and walleye pollock *Theragra chalcogramma*. Mar. Ecol.-Prog. Ser. 317, 245–258. https://doi.org/10.3354/meps317245

Witteveen, B.H., De Robertis, A., Guo, L., Wynne, K.M., 2015. Using dive behavior and active acoustics to assess prey use and partitioning by fin and humpback whales near Kodiak Island, Alaska. Mar. Mammal Sci. 31, 255–278. https://doi.org/10.1111/mms.12158

Zimmermann, M., Prescott, M.M., 2015. Smooth sheet bathymetry of the central Gulf of Alaska (NOAA Technical Memorandum No. NMFS-AFSC-287). U.S. Deptartment of Commerce.

Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. 1, 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.D., 2009. Mixed effects models and extensions in ecology with R. Springer, New York, New York, USA.

Zuur, A.F., Saveliev, A.A., Ieno, E.N., 2012. Zero inflated models and generalized linear mixed models with R. Highland Statistics Ltd., Newburgh, United Kingdom.

1049 **TABLES**

1050

1051 **Table 1 –** List of response and predictor indices used as potential predictors in station-based ("S") and
1052 transect-based ("T") models. Units in ().

| **Response variable** | **Abbreviation** | **Data Source** | **Model** |
|---|---|---|---|
| Capelin presence (binary)<br>Absent (0): $s_A = 0$<br>Present (1): $s_A > 0$ | $P$ | Derived from acoustics [1] | S, T |
| Capelin positive density ($s_A > 0$) | $D$ | Acoustics [1] | S, T |

| **Predictor variable** | **Abbreviation** | **Data Source** | **Model** |
|---|---|---|---|
| Bottom depth (m) | *BtmD* | Acoustics [1] | T |
| Bottom depth factor (binary)<br>0 = bank (< 100 m),<br>1 = trough (100 – 500 m) | *fBT* | Derived from acoustics [1] | S, T |
| Distance from 100 m depth contour (km) | *Edge* | Derived from bathymetry [2] | S, T |
| Surface temperature (° C) | *TmpS* | CTD profile [3] | S |
| Bottom temperature * (° C) | *TmpB* | CTD profile [3] | S |
| Surface-bottom temp. difference $\Delta_{TmpS\text{-}TmpB}$ (° C) | *TmpD* | Derived from CTD profile [3] | S |
| Bottom salinity * (psu) | *SalB* | CTD profile [3] | S |
| Integrated chlorophyll-α from 0-50 m (mg $m^{-2}$) | *Chla* | *In situ* bottle samples [4] | S |
| Predator: piscivorous, semi-pelagic fish ($s_A$) | *Pred* | Acoustics [1] | S, T |
| Competitor: age-0 pollock ($s_A$) | *Comp* | Acoustics [1] | S, T |
| Prey: macrozooplankton ($s_A$) | *Prey* | Acoustics [1] | S, T |

1053 * Measurement taken 10 m above bottom or 200 m, whichever was shallower
1054 [1] McGowan et al. 2016
1055 [2] Zimmermann and Prescott 2015
1056 [3] Stabeno et al., 2015
1057 [4] Strom et al., 2015
1058

**Table 2 –** Final parameter estimates (**Est)** and standard errors (**SE**) for transect-based analysis of capelin positive density relative to bottom depth (*BtmD*), season (*fSn*), year (*fYr*), and their interactions in summer (s11) and fall (f11) 2011 and summer 2013 (s13). Derived parameters for the spatial random field, including the geostatistical range (km) and the marginal standard deviation, were estimated for each survey. CV is the coefficient of variation for measurement errors. **Type** indicates if the parameter is estimated as a fixed effect (F) or is calculated as a derived quantity from other fixed effects that are not as easily interpreted (D).

| Parameter | Est | SE | Type |
|---|---|---|---|
| *B0* | 3.04 | 0.22 | F |
| *BtmD* | -1.53 | 0.27 | F |
| *fYr*:2013 | 2.49 | 0.26 | F |
| *fSn*:Fall | 0.14 | 0.33 | F |
| *BtmD*fYr*:2013 | 0.12 | 0.29 | F |
| *BtmD*fSn*:Fall | 1.96 | 0.43 | F |
| $Range_{s11}$ | 3.71 | 0.60 | D |
| $Range_{f11}$ | 5.11 | 0.98 | D |
| $Range_{s13}$ | 3.54 | 0.48 | D |
| Marginal $SD_{s11}$ | 4.28 | * 0.43 | D |
| Marginal $SD_{f11}$ | 4.17 | * 0.43 | D |
| Marginal $SD_{s13}$ | 6.30 | * 0.45 | D |
| CV | 1.00 | 0.02 | D |

**Table 3 –** Final model estimates for station-based samples from summer 2013 survey, with best models based on $AIC_C$ in bold. Model names indicate analysis resolution (s=station), response variable (P=binary presence/absence, D=positive density), and covariate combination number. **Fixed Effects** include parameter estimates (and standard errors) for intercepts and predictor indices (Table 1). **Spatial Est.** indicates derived quantities representing the geostatistical range (**Rng**, km) and marginal standard deviation (**SD**) for the spatial random field. 'X' indicates parameter included in full model, **CV** is the coefficient of variation for measurement errors in positive density models, **Delta AICc** from best model in set (AICc for best model), **Dev** is the proportion of explained variance (i.e. pseudo $R^2$).

| **Model** | ***N*** | **Fixed Effects (SE)** | | | | | | | | | | | **Spatial Est.** | | **CV** | **Delta** | **Dev** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *B0* | *fBT* | *Edge* | *TmpS* | *TmpB* | *TmpD* | *SalB* | *Chla* | *Pred* | *Comp* | *Prey* | Rng | SD | **(SE)** | **AICc** | **(%)** |
| **s.P.1** | **37** | **-1.11 (0.35)** | X | **0.56 (0.3)** | X | **0.95 (0.39)** | | | | X | X | | X | X | NA | **0.00 (45.17)** | **17.6** |
| s.P.5 | 37 | -0.45 (0.36) | -1.05 (0.62) | X | | | | X | | X | X | | X | X | NA | 2.71 | 6.7 |
| s.P.2 | 37 | -0.3 (0.46) | -1.55 (0.9) | 0.65 (0.32) | X | | | | 0.36 (0.35) | X | X | | X | X | NA | 2.78 | 17.0 |
| s.P.3 | 37 | 0.01 (0.4) | -2.18 (0.89) | 0.78 (0.35) | X | | | | | 0.31 (0.31) | X | X | X | X | NA | 2.84 | 16.9 |
| s.P.4 | 37 | -0.32 (0.55) | -1.59 (1.04) | 0.85 (0.38) | | | -0.48 (0.5) | | | 0.31 (0.32) | X | | X | X | NA | 4.5 | 19.1 |
| **s.D.1** | **12** | **5.98 (0.31)** | X | **-1.31 (0.35)** | | | | | X | X | | | X | X | **1.09 (0.19)** | **0.00 (176.32)** | **43.1** |
| **s.D.2** | **12** | **5.98 (0.31)** | X | **-1.31 (0.35)** | | | | | | X | X | | X | X | **1.09 (0.19)** | **0.00** | **43.1** |
| s.D.6 | 12 | 6.05 (0.33) | X | | | | -1.8 (0.48) | | | X | X | | X | X | 1.14 (0.2) | 1.39 | 37.3 |
| s.D.7 | 12 | 6.13 (0.34) | X | | | | | -1.17 (0.4) | | X | X | | X | X | 1.19 (0.21) | 2.81 | 30.7 |
| s.D.5 | 12 | 6.84 (0.44) | -1.89 (0.76) | | | X | | | | X | X | | X | X | 1.24 (0.21) | 4.12 | 24.2 |
| s.D.8 | 12 | 6.84 (0.44) | -1.89 (0.76) | | | | | | | X | X | X | X | X | 1.24 (0.21) | 4.12 | 24.2 |
| s.D.3 | 12 | 6.69 (0.39) | -2.06 (0.69) | | -0.91 (0.38) | | | | X | X | | | X | X | 1.1 (0.19) | 5.13 | 41.4 |
| s.D.4 | 12 | 6.69 (0.39) | -2.06 (0.69) | | -0.91 (0.38) | | | | | X | | X | X | X | 1.1 (0.19) | 5.13 | 41.4 |

**Table 4 –** Final model estimates for transect-based samples from summer 2013 survey, with best models based on $AIC_C$ in bold. Model names indicate analysis resolution (t=transect), response variable (P=binary presence/absence, D=positive density), and model combination number. **Fixed Effects** include parameter estimates (and standard errors) for intercepts and predictor indices (Table 1). **Spatial Est.** indicates derived quantities representing the geostatistical range (**Rng**, km) and marginal standard deviation (**SD**) for the spatial random field. 'X' indicates parameter included in full model, **CV** is the coefficient of variation for measurement errors in positive density models, **Delta AICc** from best model in set (AICc for best model), **Dev** is the proportion of explained variance (i.e. pseudo $R^2$).

| Model | *N* | Fixed Effects (SE) | | | | | | Spatial Est. | | CV | Delta | Dev |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *B0* | *fBT* | *Edge* | *Pred* | *Comp* | *Prey* | Rng | SD | (SE) | AICc | (%) |
| **t.P.1** | **2268** | **-8.66** | **-1.44** | **X** | **0.22** | **0.20** | **0.26** | **23.93** | **13.37** | NA | **0.00** | **71.8** |
| | | **(1.88)** | **(1.34)** | | **(0.19)** | **(0.14)** | **(0.16)** | **(3.86)** | **(2.52)** | | **(790.1)** | |
| **t.D.1** | **669** | **5.67** | **-1.79** | **-0.61** | **0.1** | **0.18** | | **5.40** | **4.13** | **0.97** | **0.00** | **79.6** |
| | | **(0.27)** | **(0.4)** | **(0.22)** | **(0.09)** | **(0.1)** | | **(0.78)** | **(0.31)** | **(0.03)** | **(8501.4)** | |
| **t.D.2** | 669 | 6.00 | -2.39 | | 0.08 | 0.20 | 0.11 | 5.68 | 4.23 | 0.97 | 5.9 | 79.6 |
| | | (0.25) | (0.37) | | (0.09) | (0.1) | (0.08) | (0.84) | (0.31) | (0.03) | | |

**Table 5 –** Final model estimates for depth-stratified transect-based samples from summer 2013 survey, with best models based on $AIC_C$ in bold. **Fixed Effects** include parameter estimates (and standard errors) for intercepts and predictor indices (Table 1). Model names indicate analysis resolution (t=transect), response variable (P=binary presence/absence, D=positive density, and depth-stratified parameters (bt). **Spatial Est.** indicates derived quantities representing the geostatistical range (**Rng**, km) and marginal standard deviation (**SD**) for the spatial random field. 'X' indicates parameter included in full model, **CV** is the coefficient of variation for measurement errors in positive density models, **AICc** for best model, **Dev** is the proportion of explained variance (i.e. pseudo $R^2$).

| Model | *N* | Fixed Effects (SE) | | | | | | | | | | Spatial Est. | | CV | | Dev |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *B0* | *fBT* | $Edge_b$ | $Edge_t$ | $Pred_b$ | $Pred_t$ | $Comp_b$ | $Comp_t$ | $Prey_b$ | $Prey_t$ | Rng | SD | (SE) | AICc | (%) |
| **t.P_bt** | **2268** | **-5.85** | **-4.26** | **2.83** | **-0.28** | **0.31** | **X** | **0.27** | **0.07** | **0.23** | **0.32** | **24.13** | **13.09** | NA | **796.1** | **71.9** |
| | | **(4.59)** | **(4.09)** | **(4.06)** | **(1.22)** | **(0.22)** | | **(0.18)** | **(0.22)** | **(0.18)** | **(0.32)** | **(3.64)** | **(2.51)** | | | |
| **t.D_bt** | **669** | **6.66** | **-3.15** | **0.77** | **-0.9** | **X** | **0.4** | **0.26** | **-0.06** | **0.12** | **0.31** | **5.27** | **4.04** | **0.96** | **8493.4** | **80.0** |
| | | **(0.54)** | **(0.64)** | **(0.69)** | **(0.23)** | | **(0.16)** | **(0.13)** | **(0.14)** | **(0.09)** | **(0.2)** | **(0.74)** | **(0.31)** | **(0.03)** | | |

**FIGURE CAPTIONS**

**Fig. 1** –Distributions of capelin acoustic density, NASC ($s_A$, $m^2$ $nm^{-2}$), in summer and fall 2011 and summer 2013, and locations of oceanography stations. Inset (red box) indicates CGOA study area. Acoustic data are categorized to indicate capelin density ("Capelin") or presence within multi-species aggregations ("Capelin present"), and "Unknown" indicates backscatter that could not be identified, "Coverage" indicates acoustic sampling, *"n*" indicates total number of 0.5 km acoustic samples (located < 500 m bottom depth) by survey, and "$s_A > 0$" indicates number of capelin positive density samples.

**Fig. 2** – Scatter plots of observed capelin acoustic positive density, $s_A$, relative to bottom depth in **(A)** summer 2011, **(B)** fall 2011, and **(C)** summer 2013; solid black lines represent predicted density across depth range of capelin occurrence, while the dashed line extends predicted density to maximum bottom depth among all surveys; red lines represent mean predicted density over banks (< 100 m, solid lines) and troughs (≥ 100 m, dotted lines). Model diagnostic plots: **(D)** observed versus predicted density, where the dotted line represents y=x relationship; **(E)** standardized sign deviance residuals versus predicted density, where red line represents a smooth spline fitted to residuals; **(F)** Q-Q plot showing sample versus theoretical quantiles of gamma distribution.

**Fig. 3** – Frequency distributions of capelin fork length from trawl samples by year, season, and bottom depth factor ("Bank" = < 100 m; "Trough" = ≥ 100 m). *n* indicates number of length samples.

**Fig. 4** – Station-based oceanographic predictor indices summarized by year and season. Scatter plots show capelin positive density, *D*, relative to each oceanographic index: "TmpS"=surface temperature (° C); "TmpB"=bottom temperature (° C); "TmpD"=surface-bottom temperature difference (° C); "SalB"=bottom salinity (psu); "Chla"=integrated chlorophyll-α (mg $m^{-2}$). Sample location is indicated by

closed circles (banks) and open circles (troughs). For summer 2013, horizontal box plots summarize each index based on presence ($s_A > 0$, dark gray) or absence ($s_A = 0$, light gray) of capelin, *P*, at the station.

**Fig. 5** – Acoustic densities of capelin and predictor indices for potential predators ("Pred", semi-pelagic piscivorous fish), competitors ("Comp", age-0 pollock), and prey ("Prey", macrozooplankton) at transect (top row) and station (bottom row) analysis resolutions from the summer 2013 survey. Inset boxplots summarize predictor indices by depth factor ("A"=all samples; "B"=bank; "T"=trough) where capelin were present (black box) and absent (gray box) for each analysis resolution. Boxplots within capelin column (far left) summarize the "Edge" predictor index, a measure of each sample's distance from the 100 m depth contour representing the edge of banks and troughs.

**Fig. 6** – Diagnostic plots for final station- and transect-based models of capelin occurrence (row **A**) and positive density (rows **B-D**) in summer 2013. Row **A**) predicted vs. observed occurrence probabilities, with predicted probabilities indicated by black line (shaded area=95% confidence interval) and mean observed probabilities by triangles. Range of predicted probabilities are in parentheses. For positive density models, row **B**) scatter plots of observed versus predicted density; dotted line represents y=x relationship; **C**) scatter plots of standardized sign deviance residuals vs. predicted density with a smooth spline fitted to the data (red line); dotted line represents y=0; **D**) Q-Q plots show sample quantiles vs. theoretical quantiles of gamma distribution. See Tables 3-5 for model name definitions (note: "Model s.D.1/2" shows final model fits for candidate models s.D.1 and s.D.2).

**Fig. 7** –Standardized parameter estimates and standard errors for fixed effects by model for capelin occurrence (upper plots) and positive density (lower plots) in summer 2013. Parameter scale indicated

by color: intercept (*BO*) and depth factor (*fBT*) are on left Y-axis in blue, environmental covariates (see Table 1 for covariate definitions) are on right Y-axis in red. "X" indicates covariates that were dropped from full model.

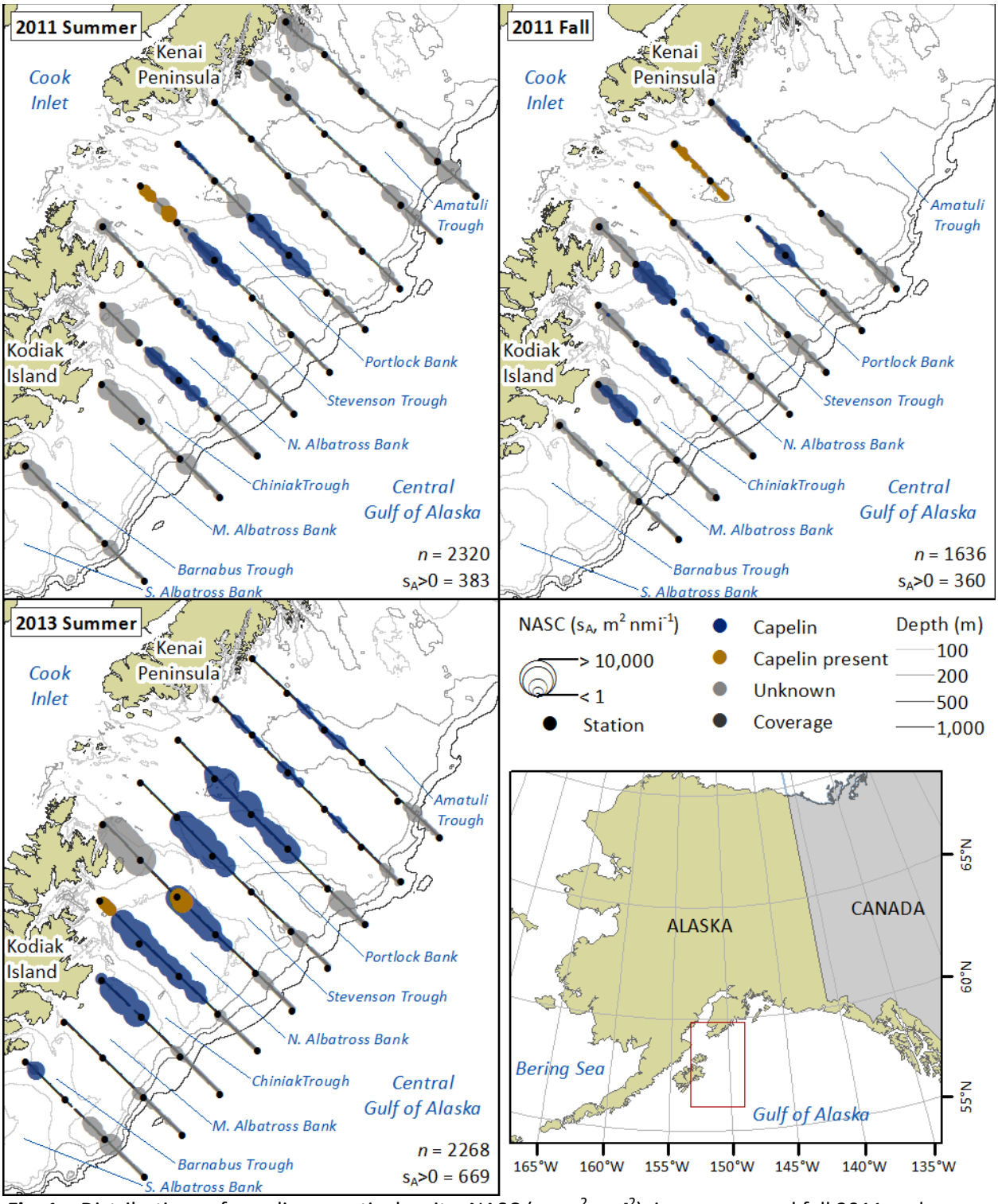


**Fig. 1** – Distributions of capelin acoustic density, NASC ($s_A$, $m^2$ $nm^{-2}$), in summer and fall 2011 and summer 2013, and locations of oceanography stations. Inset (red box) indicates CGOA study area. Acoustic data are categorized to indicate capelin density ("Capelin") or presence within multi-species aggregations ("Capelin present"), and "Unknown" indicates backscatter that could not be identified, "Coverage" indicates acoustic sampling, "*n*" indicates total number of 0.5 km acoustic samples (located < 500 m bottom depth) by survey, and "$s_A > 0$" indicates number of capelin positive density samples.

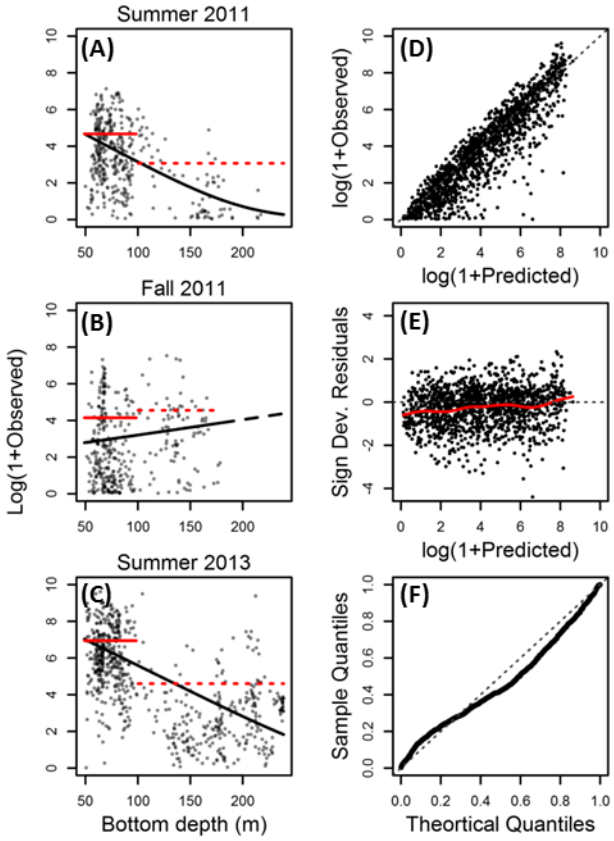


**Fig. 2** – Scatter plots of observed capelin acoustic positive density, $s_A$, relative to bottom depth in **(A)** summer 2011, **(B)** fall 2011, and **(C)** summer 2013; solid black lines represent predicted density across depth range of capelin occurrence, while the dashed line extends predicted density to maximum bottom depth among all surveys; red lines represent mean predicted density over banks (< 100 m, solid lines) and troughs (≥ 100 m, dotted lines). Model diagnostic plots: **(D)** observed versus predicted density, where dotted line represents y=x relationship; **(E)** standardized sign deviance residuals versus predicted density, where red line represents a smooth spline fitted to residuals; **(F)** Q-Q plot showing sample versus theoretical quantiles of gamma distribution.

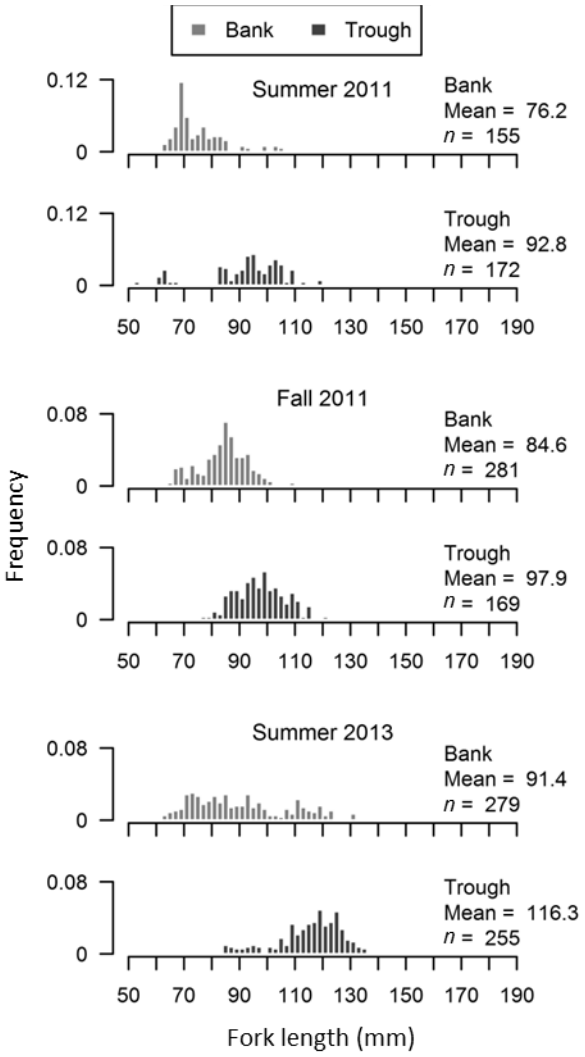


**Fig. 3** – Frequency distributions of capelin fork length from trawl samples by year, season, and bottom depth factor ("Bank" = < 100 m bottom depth; "Trough" = ≥ 100 m). *n* indicates number of length measurements.

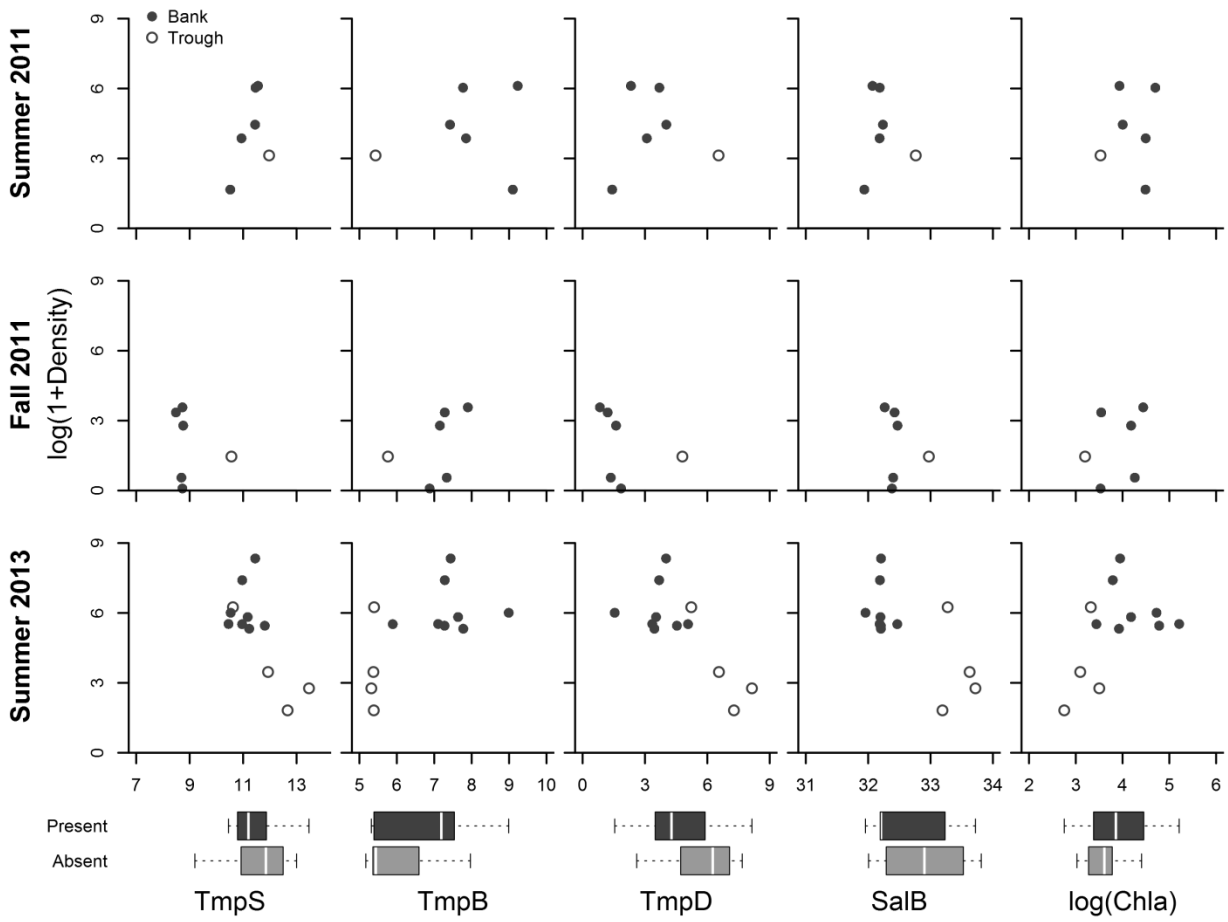


**Fig. 4** – Station-based oceanographic predictor indices summarized by year and season. Scatter plots show capelin positive density, *D*, relative to each oceanographic index: "TmpS"=surface temperature (° C); "TmpB"=bottom temperature (° C); "TmpD"=surface-bottom temperature difference (° C); "SalB"=bottom salinity (psu); "Chla"=integrated chlorophyll-α (mg $m^{-2}$). Sample location is indicated by closed circles (banks) and open circles (troughs). For summer 2013, horizontal box plots summarize each index based on presence ($s_A > 0$, dark gray) or absence ($s_A = 0$, light gray) of capelin, *P*, at the station.

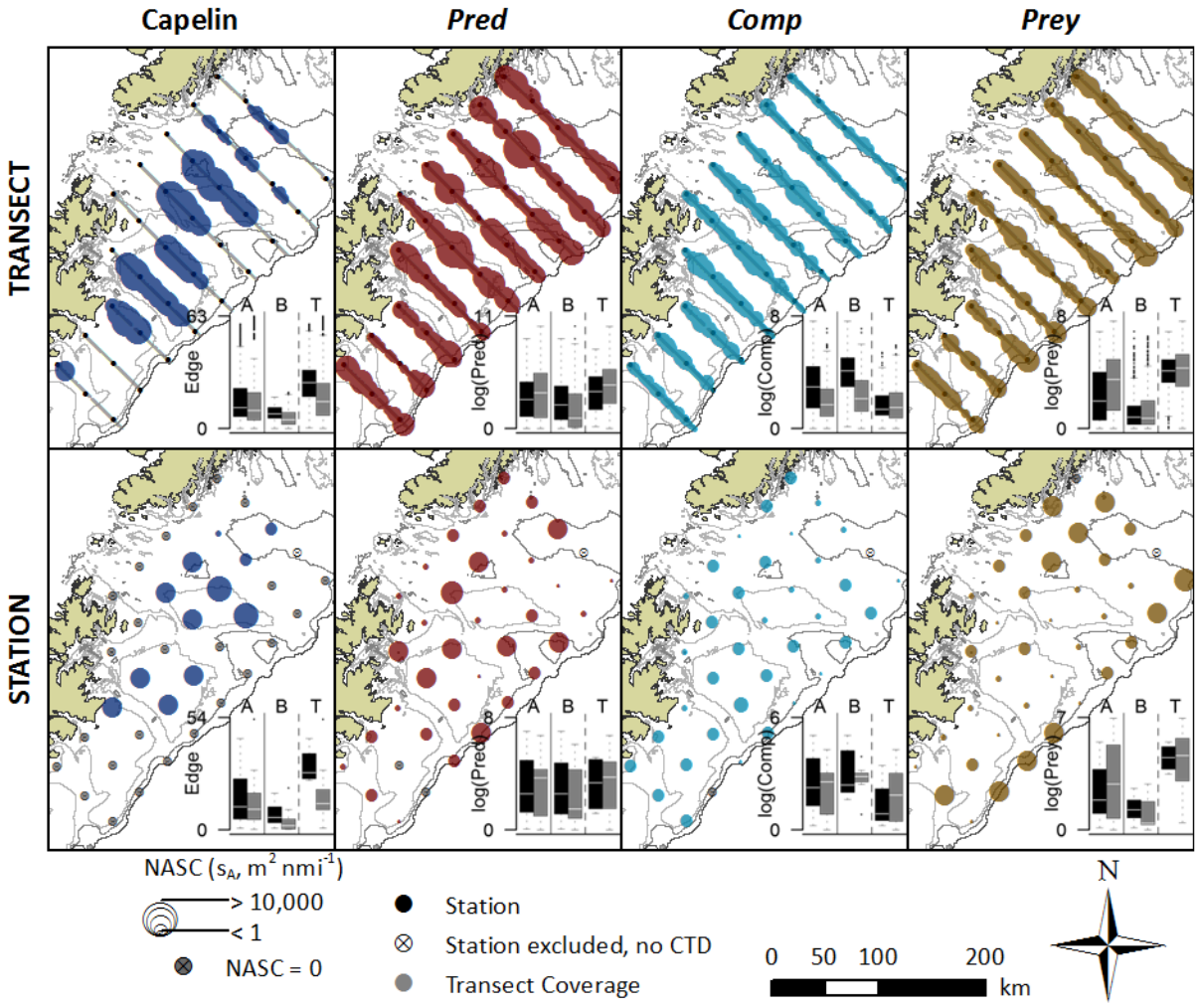


**Fig. 5** – Acoustic densities of capelin and predictor indices for potential predators ("Pred", semi-pelagic piscivorous fish), competitors ("Comp", age-0 pollock), and prey ("Prey", macrozooplankton) at transect (top row) and station (bottom row) analysis resolutions from the summer 2013 survey. Inset boxplots summarize predictor indices by depth factor ("A"=all samples; "B"=bank; "T"=trough) where capelin were present (black box) and absent (gray box) for each analysis resolution. Boxplots within capelin column (far left) summarize the "Edge" predictor index, a measure of each sample's distance from the 100 m depth contour representing the edge of banks and troughs.

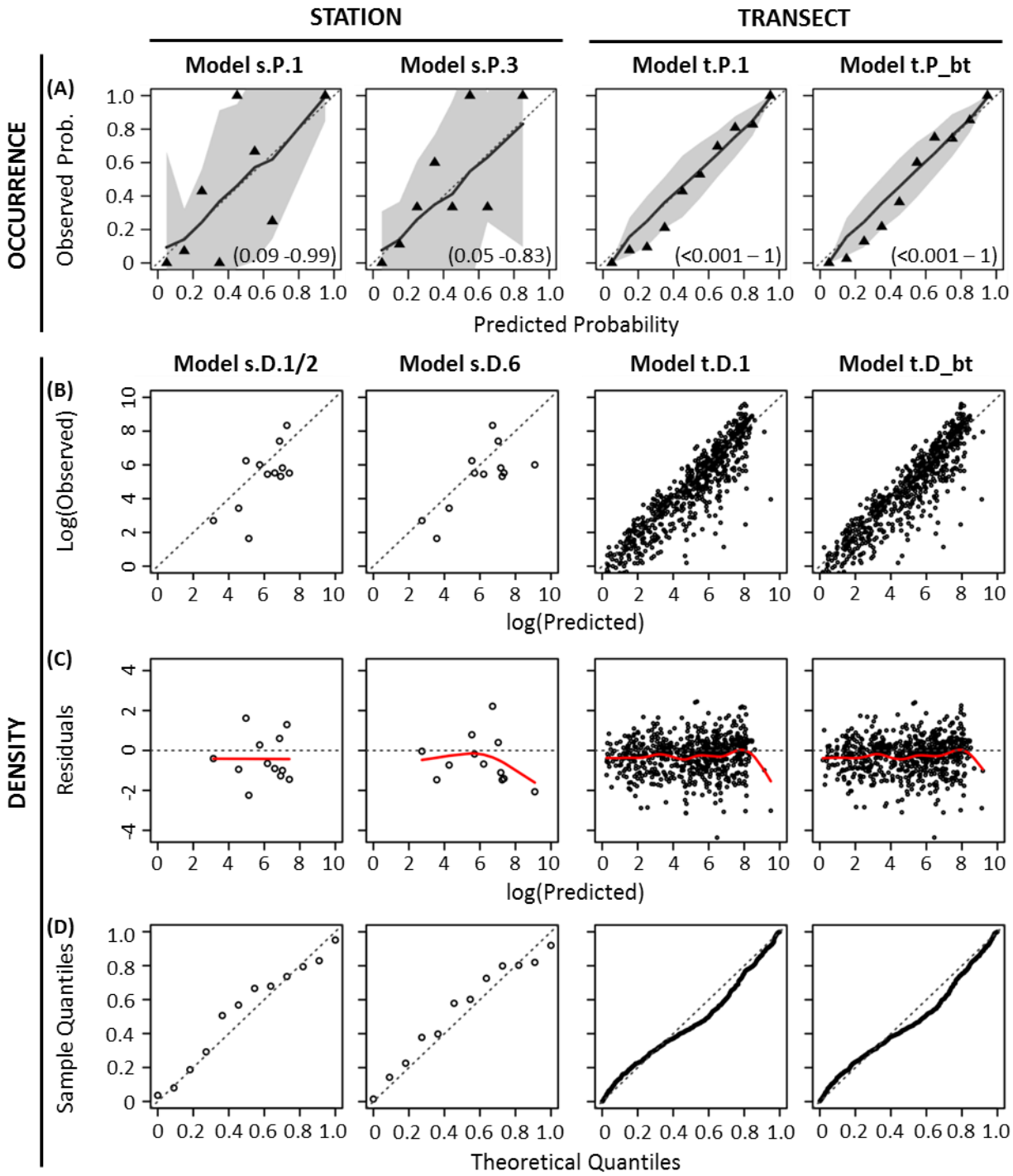


**Fig. 6** – Diagnostic plots for final station- and transect-based models of capelin occurrence (row **A**) and positive density (rows **B-D**) in summer 2013. Row **A**) predicted vs. observed occurrence probabilities, with predicted probabilities indicated by black line (shaded area=95% confidence interval) and mean observed probabilities by triangles. Range of predicted probabilities are in parentheses. For positive density models, row **B**) scatter plots of observed versus predicted density; dotted line represents y=x relationship; **C**) scatter plots of standardized sign deviance residuals vs. predicted density with a smooth spline fitted to the data (red line); dotted line represents y=0; **D**) Q-Q plots show sample quantiles vs.

theoretical quantiles of gamma distribution. See Tables 3-5 for model name definitions (note: “Model s.D.1/2” shows final model fits for candidate models s.D.1 and s.D.2).

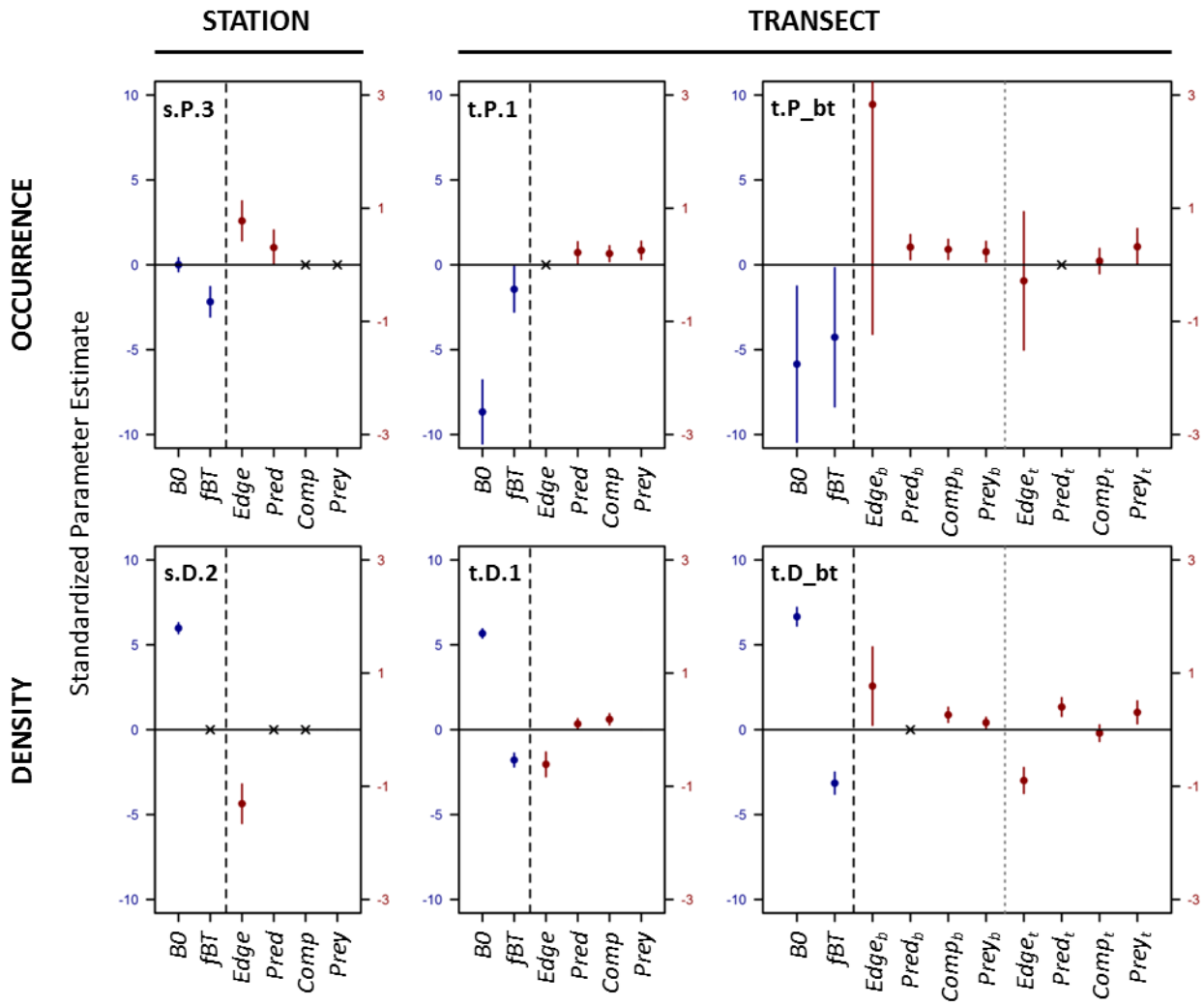


**Fig. 7** –Standardized parameter estimates and standard errors for fixed effects by model for capelin occurrence (upper plots) and positive density (lower plots) in summer 2013. Parameter scale indicated by color: intercept (*BO*) and depth factor (*fBT*) are on left Y-axis in blue, environmental covariates (see Table 1 for covariate definitions) are on right Y-axis in red. "X" indicates covariates that were dropped from full model. Note *BO* is always included and not subject to model selection using AICc.