Seasonal and interannual variation in spatio-temporal models

for index standardization and phenology studies

James T. Thorson[1], Charles F. Adams[2], Elizabeth N. Brooks[2], Lisa B. Eisner[3], David G. Kimmel[4], Christopher M. Legault[2], Lauren A. Rogers[4], Ellen M. Yasumiishi[3]

[1] Habitat and Ecological Processes Research Program, Alaska Fisheries Science Center, NOAA, Seattle, WA 98115 USA

[2] Northeast Fisheries Science Center, NOAA, NMFS, 166 Water Street, Woods Hole, Massachusetts 02543-1026 USA

[3] Ecosystem Monitoring and Assessment, Alaska Fisheries Science Center, NOAA, NMFS, Juneau, AK 99801 USA

[4] Recruitment Process Program, Alaska Fisheries Science Center, NOAA, NMFS, Seattle, Seattle, WA 98115 USA

20 **Abstract**:

21 Climate change is rapidly affecting the seasonal timing of spatial demographic processes.

22 Consequently, resource managers require information from models that simultaneously measure

23 seasonal, interannual, and spatial variation. We present a spatio-temporal model that includes

24 annual, seasonal, and spatial variation in density and then highlight two important uses: (1)

25 standardizing data that are spatially unbalanced within multiple seasons; and (2) identifying

26 inter-annual changes in seasonal timing ("phenology") of population processes. We demonstrate

27 these uses with two contrasting case studies: three bottom trawl surveys for yellowtail flounder

28 (*Limanda ferruginea*) in the Northwest Atlantic Ocean from 1985-2017, and pelagic tows for

29 copepodite stage 3+ copepod (*Calanus glacialis/marshallae*) densities in the eastern Bering Sea

30 from 1993-2016. The yellowtail analysis illustrates how data from multiple surveys can be used

31 to infer density hot spots in an area that is not sampled one or more surveys. The copepod

32 analysis assimilates seasonally unbalanced samples to estimate an annual index of the seasonal

33 timing of copepod abundance, and identifies a positive correlation between this index and cold-

34 pool extent. We conclude by discussing additional potential uses of seasonal spatio-temporal

35 models and emphasize their ability to identify climate-driven shifts in the seasonal timing of fish

36 movement and ecosystem productivity.

37

Introduction

Resource surveys are conducted worldwide to track changes in abundance for managed species. These surveys are typically designed by management agencies to support a participatory management process, where regulations are developed, proposed, and evaluated over the course of several months or years. As a consequence, resource surveys are often designed to index variation on the same interannual time-scale as management operates, typically occurring over a fixed time-interval in each year. Well-known examples include the Christmas count survey for birds in North America (Butcher *et al.*, 1990), or the spring and fall fishery resource surveys conducted annually in the Northwest Atlantic Ocean from 1963 to present day (Grosslein, 1969; Politis *et al.*, 2014). By occurring at the same time each year, this design is a powerful way to control for within-year shifts in species distribution while maximizing statistical power for detecting interannual variation in abundance.

Despite this common emphasis on capturing interannual variation in resource abundance, there are many reasons to study changes occurring within a single year. For example, many terrestrial and aquatic animals undergo a seasonal migration that partitions habitats available for feeding and raising young, or follow ontogenetic movement patterns associated with distinct life stages (e.g., immature vs. mature). The spatial extent and timing of these seasonal migrations may shift over time, either due to changes in land use (Zipkin *et al.*, 2012) or environmental conditions (Nichol *et al.*, 2019). Detecting these shifts in seasonal migration typically require data (either via individual tracks or population counts) that arise across multiple seasonal periods. In particular, the timing of different population processes ("phenology") is predicted to be a sensitive indicator of climate impacts (Scranton and Amarasekare, 2017), so detecting climate-driven shifts in seasonal processes is a high priority for climate-impact studies.

61   Furthermore, there is a growing interest in leveraging citizen-science and other opportunistic

62   data sources (Callaghan *et al.*, 2018), and proper interpretation of these data sets requires

63   accounting for variation in population density among seasons.

64   In addition to studies detecting climate impacts on phenology, there is also a growing

65   literature documenting shifts in spatial distribution that are associated with local and regional

66   climate conditions (Pinsky *et al.*, 2013). Shifts in spatial distribution can be estimated using

67   spatio-temporal models (e.g., Thorson, 2019a), which estimate a "latent" (unobserved) map of

68   population densities across space and how this map of density changes over time (Cressie and

69   Wikle, 2011). Importantly, these spatio-temporal models typically include a spatial correlation

70   function, which allows densities to be predicted even at locations where sampling does not occur

71   (Cressie *et al.*, 2009). Density predictions arising from spatially unbalanced sampling may be

72   biased when sampling locations are not "missing at random"; that is, the probability of sampling

73   is correlated with the density occurring at a given location (Cressie *et al.*, 2009; Conn *et al.*,

74   2017). Nevertheless, spatio-temporal models for spatially unbalanced data can still mitigate bias

75   arising from ignoring the spatial configuration of data, where changes in the spatial distribution

76   of sampling would otherwise be confounded with changes in population density (Thorson *et al.*,

77   2016a).

78   While scientific surveys follow a defined sampling design and aim to maintain consistent

79   protocols, there are uncontrollable circumstances that can disrupt this design. For example,

80   severe weather, mechanical problems with the ship, and constraints arising from multiple

81   objectives and/or budget cuts can lead to greatly reduced or incomplete sampling of survey strata

82   in a given year.   As a result, it becomes difficult to interpret interannual trends in abundance for

83   years when the design is not followed. In stock assessments for marine species, an analyst might

84   choose to exclude, downweight, or perform an *ad hoc* expansion for those incompletely sampled

85   years. By contrast, a seasonal spatio-temporal model could use sampling information from

86   another survey to inform about abundance in unsampled or poorly sampled areas. Similarly,

87   surveys of within-season changes in abundance or phenology will sometimes depart from their

88   design, or the design will shift over time. In these cases, it becomes necessary to simultaneously

89   correct for spatial and seasonal patterns in abundance when interpreting seasonally and spatially

90   unbalanced sampling data. While past studies have developed models that account for seasonal

91   variation when forecasting annual changes in abundance (e.g., Grieve *et al.*, 2017; Kanamori *et*

92   *al.*, 2019), these studies have not typically demonstrated the potential to assimilate seasonally

93   unbalanced sampling data (although see Pinto et al. (2018) for one exception).

94   In this study, we illustrate how to assimilate spatially and seasonally unbalanced sampling

95   data using a spatio-temporal model that includes both changes in spatial distribution among years

96   (interannual variation) and among seasons (seasonal variation). Models including both spatial

97   and temporal variation are often termed "spatio-temporal models." We start by describing a

98   default configuration of our proposed seasonal spatio-temporal model, which includes seasonal

99   and annual main effects and an autocorrelated season-year effect for both intercepts and spatial

100   variation within a delta-modelling framework. This configuration preserves correlations in

101   spatial distribution among seasons for a given year (e.g., annual drivers of distribution), among

102   years for a given season (e.g., seasonal migratory patterns), and among adjacent seasons within

103   and across years (e.g., transient hotspots in spatial distribution). Previous spatio-temporal

104   models have typically included either interannual variation (e.g., Ono *et al.*, 2018) or seasonal

105   variation (e.g., Thorson *et al.*, 2016b; Grieve *et al.*, 2017) in isolation. Alternatively, a few

106   studies have modeled residual variation in spatial distribution among seasons and years.

107 However, these have not modeled the full set of correlations we consider in this study, e.g.,

108 without correlations among years for a given season (Pinto *et al.*, 2018), without correlations

109 among seasons for a given year (Kai *et al.*, 2017), or without information about the sequence of

110 seasons within a given year (Kanamori *et al.*, 2019). We then demonstrate model performance

111 using two contrasting case-studies: (1) standardizing seasonal indices of abundance for a

112 commercially important fish in the Northwest Atlantic, and (2) estimating interannual changes in

113 phenology for a planktonic crustacean in the northeast Pacific. These two case-studies highlight

114 our ability to share information among seasons to account for spatially unbalanced sampling

115 among years and seasons, and to identify changes in seasonal dynamics over time.

Methods

116

Model structure

117

118 We seek to develop a statistical model that represents spatial variation in population density, and

119 how density changes both among years ("interannual variation") and within years ("seasonal

120 variation"). We envision that field-sampling data are available, where sample b_i records the

121 biomass (kg) or abundance (numbers) encountered by the i-th sample, occurring at location s_i

122 (within a fixed and pre-defined spatial domain), year y_i (between a minimum and maximum year

123 $y \in \{y_{min}, ..., y_{max}\}$), and season u_i (among an ordered set of intervals occurring within a year,

124 $u_i \in \{u_1, ..., u_U\}$). These seasons could be quarters (winter, spring, summer, fall), months

125 (January-December), weeks (1-52), or any intervals defined within a year (whether having even

126 or uneven spacing and duration). In this notation, the combination of year y_i and season u_i is

127 sufficient to define the time of a given sample, and the term t to describe this combination, $t \in$

128 $\{y_{min}u_1, y_{min}u_2, ..., y_{max}u_U\}$, where seasons and years are ordered such that $t + 1$ is the year-

129   season combination occurring immediately after t and $t-1$ is the year-season combination

130   preceding t. We note that densities can change rapidly within a given year and encourage future

131   studies to explore different intervals for season u, as we do in the following case study for

132   copepod densities.

133   In particular, we seek a model that has reasonable performance even when data are entirely

134   unavailable for one or more combinations of year and season. In these instances, we specify that

135   estimates in year-season t are shrunk towards predicted density in adjacent year-seasons ($t-1$

136   and $t+1$), as well as towards estimated density in other seasons for a given year (other u for a

137   given y) and density in other years for a given season (other y for a given u). This specification

138   implies that the model will include a "main effect" for season and year, as well as an

139   autocorrelated "interaction" of season and year.

140   We implement these criteria using a Poisson-link delta model (Thorson, 2018) that specifies

141   a probability distribution for random variable B, corresponding to the likelihood of response

142   variable b_i for each sample i, $\Pr(B = b_i)$. This Poisson-link delta model includes the

143   probability p_i that sample i encounters a given species (i.e., $\Pr(B > 0)$), and also the expected

144   measurement r_i given that the species is encountered, $\Pr(B|B > 0)$:

$$\Pr(B = b_i) = \begin{cases} 1 - p_i & \text{if } B = 0 \\ p_i \times g\{B|r_i, \sigma_m^2\} & \text{if } B > 0 \end{cases} \tag{1}$$

145   where we specify a lognormal distribution for the distribution g of positive catches. This

146   Poisson-link delta model predicts encounter probability p_i and positive catch rate r_i by modelling

147   two log-linked linear predictors, $\log(n_i)$ and $\log(w_i)$ for each sample i; n_i and w_i are then

148   transformed to yield p_i and r_i:

$$p_i = 1 - \exp(-a_i \times n_i) \tag{2}$$

$$r_i = \frac{a_i \times n_i}{p_i} \times w_i$$

149   where a_i is the area-swept offset for sample i. This model structure is designed such that

150   expected density d_i is the product of encounter probability and positive catch rate, and also the

151   product of transformed linear predictors (i.e., $d_i = p_i r_i = n_i w_i$). When the response-variable b_i

152   is in units biomass, these predictors can be interpreted as numbers density n_i (with units numbers

153   per area) and average weights w_i (with units biomass per number). Alternatively, if the response-

154   variable is in units numbers, n_i (with units numbers per area) and w_i (with dimensionless units)

155   describe a parametric link between expected encounter probability and expected numbers given

156   an encounter, but they are not specifically interpretable as describing numbers-density and

157   biomass per number. In both interpretations, n_i always enters via the product $a_i \times n_i$ such that

158   n_i is expressed as density.

159   The Poisson-link delta model is useful relative to other delta-models because both linear

160   predictors use a log-link function, so that all effects are additive in their impact on predicted log-

161   density. Specifically, we specify that:

$$\log(n_i) = \underbrace{\beta_n^*(t_i)}_{\substack{Year-Season \\ intercept}} + \underbrace{\omega_n^*(s_i)}_{\substack{Spatial \\ main\,effect}} + \underbrace{\xi_{nu}^*(s_i, u_i)}_{\substack{Season \\ spatial\,effect}} + \underbrace{\xi_{ny}^*(s_i, y_i)}_{\substack{Year \\ spatial\,effect}} + \underbrace{\varepsilon_n^*(s_i, t_i)}_{\substack{Year-Season \\ spatial\,effect}} \tag{3}$$

162   where the other linear predictor w_i is defined identically except that the subscript n is replaced

163   by w for all coefficients. In applications with limited information (either due to low sample sizes

164   or small variance), the variance of one or more of these terms may be estimated near zero such

165   that the corresponding term is then dropped from the model. Specifying an additive structure in

166   log-space for both variables (n_i and w_i) simplifies interpretation of estimated terms where, e.g.,

167    $\omega_n^*(s_1) = 0.1$ indicates that $n(s_1, t)$ is expected to be 10% higher at location s_1 than at location

168   s_2 where $\omega_n^*(s_2) = 0$. We also hypothesize that this additive structure in log-space will be more

169   parsimonious than a conventional delta-model, although testing this is an empirical question for

170   future research.

171   Spatial terms are estimated using a predictive-process framework (Banerjee *et al.*, 2008),

172   such that we estimate the value of each spatial variable at a set of "knots," where the number of

173   knots n_x is specified by the user in a balance between computational speed and spatial

174   resolution. In the case of spatial variation, we specify a Gaussian Markov random field for

175   vector $\boldsymbol{\omega}_n^*$ containing the value of the spatial variable $\omega_n(s)$ at each knot s:

$$\boldsymbol{\omega}_n \sim MVN(\mathbf{0}, \sigma_{n\omega}^2 \mathbf{R}_n) \tag{4}$$

176   where $\mathbf{R}_n$ is the correlation matrix and $\sigma_{n\omega}^2$ is the pointwise variance such that $\sigma_{n\omega}^2 \mathbf{R}_n$ is the

177   spatial covariance. We then project from the values $\boldsymbol{\omega}_n$ at knots to the values $\boldsymbol{\omega}_n^*$ at the location

178   of available data. The correlation matrix, in turn, is calculated based on a vector of distance

179    $\mathbf{d}(s_1, s_2)$ between any pair of locations s_1 and s_2, and we use a sparse precision matrix that

180   approximates a Matérn correlation function (Lindgren *et al.*, 2011):

$$\mathbf{R}_n(s_1, s_1) = \frac{1}{2^{\nu-1}\Gamma(\nu)} \times (\kappa_n |\mathbf{d}(s_1, s_2)\mathbf{H}|)^\nu \times K_\nu(\kappa_n |\mathbf{d}(s_1, s_2)\mathbf{H}|) \tag{5}$$

181   where we estimate a linear transformation $\mathbf{H}$ involving estimated parameters (representing

182   geometric anisotropy) and decorrelation rate κ_n. Given the value of a spatial variable at each

183   knot, the value at any given location s within spatial domain Ω is then calculated using bilinear

184   interpolation, using a projection matrix calculated by the R package R-INLA (Lindgren, 2012).

185   Season and year main spatial effects are specified similarly, except the probability of $\xi_{nu}(u)$ is

186   calculated independently for every season u, and the probability of $\xi_{ny}(y)$ is calculated

187   independently for every year y. However, the year-season interaction $\boldsymbol{\varepsilon}_n(s,t)$ is autocorrelated

188   across the ordered year-season index t:

$$\boldsymbol{\varepsilon}_n(s,t) \sim \begin{cases} MVN(\mathbf{0}, \mathbf{Q}^{-1}) & \text{if } t = t_{min} \\ MVN(\rho_{n\varepsilon}\boldsymbol{\varepsilon}_n(s, t-1), \mathbf{Q}^{-1}) & \text{if } t > t_{min} \end{cases} \tag{6}$$

189   where the degree of autocorrelation $\rho_{n\varepsilon}$ in spatio-temporal variation $\boldsymbol{\varepsilon}_n(s,t)$ is also estimated.

190   We here assume that the decorrelation-rate parameters κ_n and κ_w are identical for spatial and

191   spatio-temporal components and different between the two linear predictors n and w; future

192   research could explore alternative specification for these hyperparameters.

193   Similarly, intercepts $\beta_n(t)$ are specified such that they can be interpolated for season-year

194   combinations without any data using information from adjacent season-years, other years of the

195   same season, or other seasons of the same year. This is again accomplished by including season

196   and year main effects, and an autocorrelated interaction of season and year:

$$\beta_n^*(t) = \mu_\beta + \beta_{nu}(u) + \beta_{ny}(y) + \beta_{nt}(t) \tag{7}$$

197   Where μ_β is the average intercept across all seasons and years, $\beta_{nu}(u)$ captures differences in

198   expected intercept among seasons u, $\beta_{ny}(y)$ captures differences in expected intercepts among

199   years y, and $\beta_{nt}(t)$ represents an autocorrelated season-year interaction:

$$\beta_{nt}(t) \sim \begin{cases} N(0, \sigma_{n\beta}^2) & \text{if } t = t_{min} \\ N(\rho_{n\beta}\beta_{nt}(t-1), \sigma_{n\beta}^2) & \text{if } t > t_{min} \end{cases} \tag{8}$$

200   and where the magnitude of autocorrelation is again estimated from available data. We ensure

201   identifiability for μ_β, $\beta_{nu}(u)$, and $\beta_{ny}(y)$ by imposing a corner constraint on the season and year

202 effects (i.e., $\beta_{nu}(u) = 0$ and $\beta_{ny}(y) = 0$ for the first season u and year y). This corner

203 constraint is necessary for intercepts (i.e., Eq. 7) but not spatial terms (i.e., Eq. 3) because the

204 season and year intercepts are treated as fixed effects, while the season and year spatial terms are

205 treated as random effects. We note that this model structure imposes no constraints on the

206 expected "shape" of seasonal variation; that is, the model can capture unimodal or multi-modal

207 distribution of abundance across seasons within a year.

208 Parameters are estimated using release 3.2.0 of package VAST (Thorson, 2019b), which is

209 publicly available online (https://github.com/James-Thorson/VAST), and runs within the R

210 statistical environment (R Core Team, 2017). The performance of VAST for models such as this

211 has been simulation-tested previously (Thorson *et al.*, 2015, 2016a; Grüss *et al.*, 2019; Johnson

212 *et al.*, 2019; Brodie *et al.*, 2020), and we recommend future simulation experiments exploring

213 performance for the specific seasonal structure proposed here. VAST estimates fixed effects

214 while approximating their marginal likelihood using the Laplace approximation (Skaug and

215 Fournier, 2006). The Laplace approximation is implemented in turn using R package TMB

216 (Kristensen *et al.*, 2016), and computational efficiency is improved using automatic

217 differentiation (Fournier *et al.*, 2012) and the SPDE approximation to spatial correlation matrices

218 (and associated projection matrices) from R-INLA (Lindgren and Rue, 2015). Standard errors

219 are calculated using a generalization of the delta-method (Kass and Steffey, 1989), and standard

220 errors are available for predictions of local density after estimating all fixed and random effects:

$$d(s,t) = n(s,t) \times w(s,t) \tag{9}$$

$$= \exp\{\beta_n^*(t) + \omega_n^*(s) + \xi_{nu}^*(s,u) + \xi_{ny}^*(s,y) + \varepsilon_n^*(s,t)\}$$

$$\times \exp\{\beta_w^*(t) + \omega_w^*(s) + \xi_{wu}^*(s,u) + \xi_{wy}^*(s,y) + \varepsilon_w^*(s,t)\}$$

221   where density can then be visualized or further processed to yield derived statistics. For

222   example, total abundance is calculated as the area-weighted sum of density $d(s,t)$ predicted at a

223   fine spatial resolution:

$$D(t) = \sum_{s=1}^{n_s} a(s)d(s,t) \tag{10}$$

224   Where n_s is the number of fine-scale predictions and $a(s)$ is the spatial area associated with each

225   prediction.

226   **Case study demonstrations**

227   We apply this model to two case studies that are chosen to represent different geographic areas,

228   seasonal resolution, taxonomic focus, and inferential goals.

229   1. *Yellowtail flounder on Georges Bank*: The first case-study involves yellowtail flounder

230   (*Limanda ferruginea*) on Georges Bank, a shallow stretch of the continental shelf in the

231   Northwest Atlantic Ocean. This commercially important species is sampled by three bottom

232   trawl surveys that each provide a different snapshot of fish distribution and abundance. We

233   seek to answer: can we standardize data from multiple surveys to share information across

234   seasonal sampling programs and thereby predict density in unsampled times and areas?

235   2. *Copepods in the eastern Bering Sea*: The second case-study involves copepodite stage 3-6

236   abundance in oblique plankton tows from five meters above sea floor to water surface for

237   *Calanus marshallae/glacialis*. These copepods are a substantial component of diet for fish,

238   seabirds and marine mammals in the eastern Bering Sea (Livingston *et al.*, 2017). Copepods

239   undergo rapid population growth and decline during the annual sampling interval, and this

240   application is therefore useful to highlight the potential for a relatively fine resolution for

241   seasonal intervals. We seek to answer: can seasonally unbalanced sampling data be used to

242   identify changes in the seasonal timing of population densities; for example, to estimate the

243   seasonal match between larval fish and zooplankton prey?

244   We discuss each in more detail below (see Fig. 1 for seasonal and interannual variability in data

245   availability).

246   *Case study #1: Yellowtail flounder on Georges Bank*

247   Yellowtail flounder are a historically important component of the Georges Bank fishery and

248   ecosystem (Stone *et al.*, 2004), but are currently at low biomass due to low recruitment and

249   unexplained mortality of adults (Legault and McCurdy, 2018). Despite a large amount of age-

250   based data available for this stock, standard stock assessment models have performed poorly

251   with large retrospective patterns and were replaced with an index-based approach in 2014

252   (Legault *et al.*, 2014).

253   We compile sampling data for three bottom trawl surveys 1985-2017 (see Fig. S1): Spring

254   (March 24 – April 25) and Fall (Sept. 30 – Nov. 6) bottom trawl surveys conducted by the

255   Northeast Fisheries Science Center (NEFSC), and a bottom trawl survey by Department of

256   Fisheries and Oceans (DFO) Canada (Feb. 11 – March 29) (O'Boyle *et al.*, 1995; Politis *et al.*,

257   2014). The NEFSC Spring and Fall surveys switched vessels in 2009, and we use data converted

258   to units of the earlier survey based on paired intercalibration sampling (Miller *et al.*, 2010). By

259   contrast, the DFO survey gear has been consistent over time, and uses different protocols than

260   the earlier US survey, hence likely catching a different proportion of local abundance (termed

261   "catchability") than the US surveys. Surveys also differ in their spatial footprint of sampling: the

262   US surveys follow a stratified random design that covers the entire stock boundary, while the

263 DFO survey uses a different stratified random design and samples only a subset of this spatial

264 area, avoiding the southwestern corner as well as moraine habitat in the western portion (Stone

265 and Gross, 2012).

266 Design-based indices of abundance are available for all three surveys, although we note that

267 the DFO design-based index is calculated only for the portion of the stock sampled by that

268 survey. We also fit an alternative multivariate spatio-temporal model, which treats each season

269 as a separate category and estimates independent spatial and spatio-temporal variation for each

270 season. This "default" specification is widely used in fisheries index-standardization (Thorson,

271 2019b), and therefore provides a reference-value for abundance indices calculated in a more

272 conventional manner.

273 The Georges Bank yellowtail flounder stock is defined for an area that straddles the Hague

274 Line separating the US and Canadian exclusive economic zones (EEZ), requiring a method to

275 allocate the total quota between the two countries. Swept area estimates from the three surveys

276 (DFO, NEFSC spring and NEFSC fall) are averaged, proportions on either side of the Hague

277 Line are calculated, and then a loess smooth is fit to the respective time series east/west. The

278 total annual quota is allocated between countries based on the terminal year smoothed proportion

279 of total stock biomass in US and Canadian waters (Murawski and Gavaris, 2004). Given the

280 importance of comparing estimates of abundance between DFO and US surveys, we seek to

281 develop an alternative model that can predict density using data from all surveys over the spatial

282 footprint of the stock assessment.

283 We use a model with annual and seasonal main effects for both intercepts and spatial

284 variation, using 250 knots to approximate spatial variables, and an autoregressive process for the

285 interaction of year and season. Initial runs of the full model showed that several variance

286　components were estimated to be zero, and we changed the model to avoid numerical problems

287　associated with a parameter estimated at a bound (i.e., an estimated variance of zero); see Table

288　S3 for details. The seasonal main effect for intercepts includes the effect of both seasonal

289　differences in density as well as differences in catchability among gears, and resulting density

290　estimates are not directly comparable among surveys due to this missing information regarding

291　catchability differences. We therefore display resulting abundance indices after scaling each

292　series to have a mean of 1.0 across years. We focus interpretation by displaying the

293　autocorrelated spatial variation in the season-year interaction, to demonstrate how this model

294　component allows information to be shared between seasons.

295　*Case study #2: Copepods in the eastern Bering Sea*

296　Copepods are an important component of secondary production in the Bering Sea, and occur in

297　stomach samples for a wide range of commercially important fishes (Livingston *et al.*, 2017).

298　Two species of copepod (*Calanus glacialis* and *C. marshallae*) co-occur in the eastern Bering

299　Sea and are difficult to identify to species, and we refer to them collectively as *C.*

300　*glacialis/marshallae* in the following (Campbell *et al.*, 2016). *C. glacialis/marshallae* has one

301　cohort per year in the eastern Bering Sea and is a key prey for multiple fish species in their first

302　year of life (Strasburger *et al.*, 2014). Productivity and spatial dynamics in this ecosystem are

303　strongly associated with wintertime sea-ice extent (Hunt *et al.*, 2011) and the resulting extent of

304　cold near-bottom waters (termed the 'cold pool;' Sigler *et al.*, 2011) during the summer growing

305　season for copepods (Eisner *et al.*, 2014, 2018; Kimmel *et al.*, 2018). Historically, zooplankton

306　sampling in the EBS occurred on process-oriented surveys, which varied year to year in their

307　spatial and temporal coverage. Recently, sampling has occurred more regularly in spring (late

308　April) and fall (late September) along the 70 m isobath from the southeastern Bering Sea M2

309   mooring location (56.87N, 164.05W) to the M8 mooring location (62.19N, 174.69W).

310   Additionally, zooplankton sampling has occurred in recent years along a grid in the southeastern

311   Bering Sea from the inner to outer shelf (~ 30-180 m) from 54.5 to 60 N in May and again in

312   August/September. We here restrict samples to oblique tows of a bongo net, starting near bottom

313   and ending near surface (see Fig. S2), and approximate spatial variation using 100 knots. These

314   oblique tows include a measure of total volume swept, and we divide the total number of

315   copepodite stage 3+ copepod (*Calanus glacialis/marshallae*) by volume swept, and then

316   multiply by the seafloor depth at the beginning of the tow to obtain vertically integrated

317   numbers-density. Using vertically integrated numbers-density as response-variable then allows

318   us to predict vertically integrated densities across a standard survey area, where the sum across

319   this survey area represents a prediction of vertical and spatially integrated abundance in numbers.

320   Future studies could instead include water-column height as an offset, and expand densities

321   while defining area, $a(s)$ in Eq. 10, as the product of surface-area and water-column height; we

322   leave this comparison as a topic of future research. *C. glacialis/marshallae* occur in 97% (3691

323   of 3802) of all available samples, and there is little information available to identify spatial,

324   temporal, or seasonal patterns in encounter probability. We therefore use a reduced model

325   wherein we remove all spatial and temporal variation in the first linear predictor (i.e., $n_i = n$ for

326   all locations s, seasons u, and years y), such that the model estimates a single encounter

327   probability for all samples, and focuses inference on the expected number of *Calanus*

328   encountered in each sample. The model again estimated several variance terms near zero, and

329   we eliminated terms with zero variance to avoid numerical difficulties (see Table S3 for details).

330   We specifically seek to measure interannual changes in *C. glacialis/marshallae* phenology.

331   To do so, we calculate the weighted average of seasonal intervals, where each season is weighted

332   by the total abundance in that season. This "abundance-weighted average season" statistic is

333   analogous to "center of gravity" (COG), which is calculated as the abundance-weighted spatial

334   location for a population; COG is widely used as a measure of spatial distribution shifts (Pinsky

335   and Palumbi, 2013; Thorson *et al.*, 2016a). We specifically calculate abundance-weighted

336   average season $\bar{u}(t)$ as:

$$\bar{u}(y) = \sum_{t \in \mathbf{t}_y} \left(\frac{D(t)}{D_y} \times u_t \right) \tag{11}$$

337   where $\mathbf{t}_y$ is the vector of season-year indices t in year y, u_t is the season corresponding to

338   season-year t, D_y is the area-weighted sum of densities across locations and seasons in year y,

339     $D_y = \sum_{t \in \mathbf{t}_y} D(t)$, and $D(t)$ is defined in Eq. 10. Standard errors for $\bar{u}(t)$ are again calculated

340   using the generalized delta-method. This index does not measure "peak abundance" *per se*, but

341   instead measures years when *C. glacialis/marshallae* abundance is high in early or later portions

342   of the summer growing season. Given the known climate-sensitivity of copepod production in

343   the Bering Sea, we compare our phenology estimate with the summertime areal extent of the

344   cold pool (waters < 2 °C) on the EBS shelf (Robert Lauth, personal communication) as measured

345   using temperature sensors deployed on bottom trawl gear in the eastern Bering Sea (Lauth and

346   Conner, 2016).

347   Finally, we explore the sensitivity of the abundance-weighted average season $\bar{u}(t)$ to the *a*

348   *priori* specification of the number of seasonal intervals that are modeled in each year. To do so,

349   we contrast results from our model using month-long intervals (nine seasonal-intervals per year

350   after excluding months with no sampling) with either aggregating all samples into Spring (Feb.-

351   May) and Summer/Fall (June-Oct.) seasons (two seasons per year), or aggregating samples into

352   two-week intervals (18 seasonal intervals per year with data). For each of these seasonal

353   specifications, we fit the same seasonal spatio-temporal model and compare resulting estimates

354   of $\bar{u}(t)$; we hypothesize that either month-long or biweekly intervals will be sufficiently fine-

355   grained to yield similar results, while the two-season model will have such coarse seasonal

356   resolution that it will likely result in different estimates than the other two models.

Results

358   We use the same seasonal spatio-temporal model for both case studies (see Fig. S3 for

359   triangulated meshes in each example). However, we use this model for very different inference

360   in each case study, and therefore address them separately below.

Georges Bank yellowtail flounder: Inferring seasonal variation in unsampled habitat

362   We first compare total abundance of yellowtail flounder on Georges Bank in each of three

363   seasons, estimated using the seasonal spatio-temporal model with either a design-based index or

364   a simpler spatio-temporal model that is run independently for each season (Fig. 2). This

365   comparison shows that the two model-based indices have similar estimates of total abundance,

366   and differ more from the design-based index for DFO. We note that the design-based index for

367   DFO has a large spike in 2008/2009, and this spike is not represented in either model-based

368   index. Model exploration shows that this spike arose from one sample in each of those years

369   wherein the DFO survey caught an anomalously large quantity of yellowtail flounder. The

370   model-based indices both treat these observations as "outliers" relative to the weighting used by

371   the design-based estimator, and hence these samples have less statistical leverage in the model-

372   based than the design-based index.

18

373   We next explore the ways in which the seasonal spatio-temporal model shares information

374   among seasons. The largest consequence is that it extrapolates the DFO index back to

375   1985/1986, when there was no DFO sampling (see Fig. 2 top panel and Fig. S1). This

376   extrapolation is based on the main effects of season and year for both intercepts and spatial

377   variation, such that the index in those years is "shrunk" towards the approximately stable trend in

378   abundance seen in the other seasons in that year; the extrapolation also results in an increase in

379   standard errors (width of blue shaded area) in those years relative to others. We also highlight

380   the consequence of autocorrelation ($\rho_{n\varepsilon} = 0.85$ and $\rho_{w\varepsilon} = -0.19$; see Table 1) on estimated

381   spatio-temporal variation arising across seasons and years (Fig. 3). In 1989-1992, for example,

382   the spring and fall surveys both encountered an increase in density relative to long-term averages

383   in the southwestern stock area, where DFO sampling does not occur. This hot spot is estimated

384   to occur starting in fall 1989, and persists consecutively through 1990 and into the DFO survey

385   in 1991, before largely reverting to long-term densities by the spring survey in 1991. These and

386   other examples of autocorrelated season-year variation cause the seasonal model to shrink

387   estimates in areas lacking samples towards their estimates in other seasons with more spatially

388   comprehensive sampling.

389   **Eastern Bering Sea copepods: Measuring interannual variation in phenology**

390   We next explore the behavior of the seasonal model when applied to a case study with more

391   seasons (nine intervals each lasting one month). This case-study exhibits an extremely

392   unbalanced sampling design across seasons (see bottom panel of Fig. 1), where the majority of

393   sampling is available in either April-May or Aug.-Sept. We start by illustrating *C.*

394   *glacialis/marshallae* density estimates in four selected years (Fig. 4, showing

395   1994/2002/2007/2012). Later years (2007/2012) are selected to show the impact of the spatially

396   distributed sampling design in May and Aug.-Oct. that is currently implemented in the eastern

397   Bering Sea, while earlier years (1994/2002) are selected to show the relatively opportunistic

398   sampling that occurred in these years (see Fig. S2 for spatial coverage of sampling across all

399   years). Density plots (Fig. 4) again show the role of autocorrelated season-year effects ($\rho_{w\varepsilon} =$

400   0.72; see Table 1) on model performance where, for example, the spatial configuration of high-

401   density areas in 2012 is essentially interpolated between May and August, given that sampling

402   data are sparse between these two months. Alternatively, years with few sampling data in any

403   month (e.g., 1994) are shrunk towards the estimate of the season main-effect, where the model

404   typically estimates increased densities and a broader distribution of high-density habitats in June-

405   July relative to early (Feb.-April) or late (Sept.-Oct.) months.

406   We next illustrate abundance indices across seasons for each year (Fig. 5), which confirms

407   that the seasonal main-effects are estimated to follow a dome-shaped pattern in *C.*

408   *glacialis/marshallae* density with a peak in June-July (as shown in 1996/2001, when data are

409   particularly sparse). As noted in the Methods section, this dome-shaped pattern is not specified

410   within the estimation model, and instead arises purely from average patterns seen in the data.

411   Similarly, the action of the year main-effect is also apparent. For example, sampling in April

412   1994 yields elevated density relative to its spatial and seasonal expectation and the model

413   therefore predicts elevated density for all seasons in that year. By contrast, sampling in August

414   2002 yielded lower densities than the average for its season and location, resulting in an estimate

415   of low density for that entire year. In these years with seasonally restricted sampling (e.g.,

416   1993/2002), the confidence intervals are relatively wide compared with years with seasonally

417   distributed sampling (e.g., 2007-2012).

418   Finally, we compute the abundance-weighted average season for *C. glacialis/marshallae* in

419   the eastern Bering Sea (Fig. 6). Standard errors are sufficiently small to detect interannual

420   variation in phenology (e.g., where 2003 is earlier than 2007) during years with seasonally

421   distributed sampling (approximately 2003 onward), while intervals are much wider in earlier

422   years. The predicted index across all years ranges between mid-June to late July, and therefore

423   spans nearly 45 days. As expected, the index is positively correlated (Pearson correlation = 0.60)

424   with cold-pool extent, where "warm years" (those with a small cold pool) are estimated to have

425   earlier average copepod phenology and vice-versa for "cold years" with a large cold pool. The

426   index is similar when specifying 18 biweekly seasonal intervals (Fig. S4). Indices from these

427   two models are also correlated with the index estimated using two (Spring vs. Summer/Fall)

428   seasonal intervals, although the latter model does not yield an interpretable scale for variation

429   among years due to the coarse resolution of seasonal intervals.

Discussion

431   In this study, we have demonstrated how a seasonal spatio-temporal model can be constructed to

432   account for spatially and seasonally unbalanced sampling data. We used two contrasting case

433   studies to show that this seasonal spatio-temporal model can yield information that is useful for

434   both ecological insight as well as applied fisheries management. We first discuss each of these

435   case studies in detail, and then speculate about other potential applications.

436   Our first case study focused on inferring interannual changes in resource abundance for a

437   commercially important fish while accounting for spatial areas that are regularly sampled by

438   some but not all seasonal surveys. This situation is common in many transboundary

439   negotiations, for example, for fisheries in the North Sea (Pedersen and Berg, 2017), although the

440   yellowtail flounder case-study had a unique difficulty of involving different spatial coverage for

441   data in each season. While both the NEFSC surveys and the DFO survey have a stratified

442   random design, the DFO survey has a higher sampling rate in Canadian waters than US waters

443   (see Fig. S1). In particular, the portion of the DFO survey used in the Georges Bank yellowtail

444   flounder assessment does not cover the southwest portion of the bank, which is covered by the

445   two NEFSC surveys. There are DFO survey strata in this area, but this area often does not

446   contain sample locations due to the low sampling rate or is dropped entirely when the survey has

447   to end early due to weather or vessel troubles. These practical and logistical factors are common

448   in bottom trawl surveys and lead to imbalanced sampling and coverage among surveys. Applying

449   a seasonal spatio-temporal model provides a statistically sound approach to combining the

450   information from the multiple surveys; we showed that it is useful specifically to infer hot spots

451   in density in unsampled areas and times based on sampling occurring in other seasons. This may

452   be particularly important when the spatial distribution of biomass is estimated for areas that

453   differ from the survey stratification, as in the quota allocation sharing agreement for this stock.

454   By contrast, the second case study focused on inferring interannual variation in phenology

455   for two pelagic crustacean species using vertically integrated tows from many different sampling

456   programs, each with different seasonal timing and spatial distribution. Resulting estimates of

457   abundance-weighted average season showed a clear and statistically significant association

458   between phenology and oceanographic conditions, while allowing comparison of phenology

459   across years with greatly different seasonal data availability. Numerous studies have shown that

460   *C. glacialis/marshallae* populations have peak abundance in June and July (e.g., Eisner et al.

461   (2018) using data from 2008-2010), and also respond to interannual variability in temperature on

462   the southeastern Bering Sea shelf (Baier and Napp, 2003; Coyle *et al.*, 2008; Eisner *et al.*, 2014,

463 2018; Coyle and Gibson, 2017; Kimmel *et al.*, 2018). In years with early ice retreat, *C.*

464 *glacialis/marshallae* that emerge from diapause will experience a potential mismatch with the

465 spring phytoplankton bloom that occurs later during warm years (Brown and Arrigo, 2013;

466 Sigler *et al.*, 2014, 2016). This mismatch combined with warmer temperatures that accelerate

467 development, appears to reduce *C. marshallae/glacialis* abundances and they typically have low

468 abundance by mid-August to late September in warm years, perhaps due to entry into diapause

469 (Kimmel *et al.*, 2018). During years of late ice retreat, by contrast, *C. glacialis/marshallae*

470 benefits from the presence of ice-associated algae that are grazed heavily to fuel egg production

471 (Campbell *et al.*, 2016). Cooler temperatures slow development rates and allow accumulation of

472 individuals that persist on the shelf due to a delay of entry into diapause (Coyle and Gibson,

473 2017; Kimmel *et al.*, 2018). Continued warming of ocean waters is expected to impact

474 phenology of *Calanus* by shifting the timing of emergence and exit into diapause (Wilson *et al.*,

475 2016). Long-term warming has already caused earlier appearance of *C. glacialis* in the White

476 Sea (Persson *et al.*, 2012; Usov *et al.*, 2013) and *C. finmarchicus* in Svalbard and northern

477 Iceland (Espinasse *et al.*, 2018).

478 Improved information regarding seasonal timing derived from seasonal spatio-temporal

479 models could be useful for a wide variety of ecosystem studies and management questions. In

480 the eastern Bering Sea, for example, the abundance of *C. marshallae/glacialis* during the fall is

481 associated with walleye pollock (*Gadus chalcogrammus*) condition and early-life survival

482 (Heintz *et al.*, 2013; Siddon *et al.*, 2013; Eisner and Yasumiishi, 2018). Notably, positive

483 significant linear relationships have been found between late summer abundances of large

484 copepods, and abundances of age-3 pollock three years later when they enter the fishery (Eisner

485 and Yasumiishi, 2018). By estimating how phenology of *C. glacialis/marshallae* varies from

486   year to year, this study shows that changes in phenology likely contribute to interannual variation

487   in late summer abundance. Detecting changes in phenology can affect the seasonal timing of

488   spawning migrations and therefore the interpretation of pre-season test fisheries (Flynn and

489   Hilborn, 2004).

490   In addition to being useful for detecting changes in phenology, the approach described here

491   could be useful for standardizing survey indices that are potentially biased by seasonal changes

492   in abundance or distribution. For instance, for taxa, stages, or processes that develop rapidly in

493   time and vary over space (e.g., zooplankton, larval fish, and squid abundance) small changes in

494   survey timing can result in biased estimates of abundance. Seasonal spatial-temporal models

495   could be used to control for survey timing and extent. In a similar vein, fisheries scientists

496   continue to use catch-per-unit-effort from fishery data to index changes in fish abundance.

497   Previous research has shown how spatio-temporal models can be used to control for interannual

498   variation in the spatial distribution of fishing (Thorson *et al.*, 2017; Grüss *et al.*, 2019), but there

499   has been less focus on controlling for differences in spatial distribution occurring at finer

500   seasonal scales than quarterly (i.e., >4 seasons per year). Seasonal fishery data could be

501   particularly useful in conjunction with resource surveys, for example, where a seasonal spatio-

502   temporal model could be fitted to both data sources to index changes in phenology that affect the

503   proportion of fish that are available to a given survey (Nichol *et al.*, 2019). In addition, such an

504   approach could be used to study how seasonal shifts in distribution are changing in response to

505   climate, which may in turn affect availability to resource users or be useful as an environmental

506   index within a stock-assessment model.

507   Results are complementary to the growing worldwide effort to document changes in the

508   spatial distribution of fishes and many other taxa (Anderson *et al.*, 2009). In fishes, the rate of

509 climate-driven shift in distribution is often measured using "center of gravity," ; that is, the

510 centroid of the distribution measured as latitude, longitude, depth, or other location measures

511 (Pinsky *et al.*, 2013; Thorson *et al.*, 2016a; Currie *et al.*, 2019). Center of gravity is a useful

512 measure for comparing rates of climate-driven distribution shifts across regions but does not

513 capture specifics about the density available to individual fishing ports or other partitions of

514 available habitat (Rogers *et al.*, 2019; Selden *et al.*, 2020). Similarly, we use abundance-

515 weighted average season as a measure of seasonal timing for available copepod prey; this

516 follows similar practice measuring changes in phenology using a population average (Rogers and

517 Dougherty, 2019). While this aggregate measure of timing does not indicate total abundance or

518 availability in specific seasons, we believe that this metric will provide a useful basis for

519 comparisons of climate-driven shifts in phenology among regions, and hope that having a

520 standardized measure facilitates future comparative research.

521   We note that seasonal spatio-temporal models also generate a new set of interesting questions

522 to resolve when using results. For example:

523 1. Is the spring, fall, or DFO index of yellowtail flounder abundance most appropriate to

524   include when fitting a stock assessment model, or perhaps some average of these different

525   surveys?

526 2. Is late summer/fall abundance of *C. glacialis/marshallae* the best predictor of pollock

527   recruitment, or instead the abundance of *C. glacialis/marshallae* across all seasons within a

528   given year (the area under the curve in each panel of Fig. 5)?

529 3. Is abundance-weighted average season the best metric for measuring changes in phenology,

530   or is it better to measure changes in the seasonal timing of emergence/disappearance of

531   copepods, analogous to measuring leading/trailing edges in spatial measures of distribution

532   (Anderson *et al.*, 2009)?

533   These and other questions become apparent when explicitly accounting for seasonal variation

534   within spatio-temporal models, and the solution will obviously depend upon how results are

535   being used. Thus, these models open up new opportunities for study, as well as challenges for

536   interpretation. We therefore argue that analysts should communicate clearly about how they

537   interpret seasonal variation in any future applications of the seasonal spatio-temporal model. We

538   also note that developing sensitive diagnostics for model mis-specification, and simulation-

539   testing the likely performance of spatio-temporal models is an ongoing endeavor. The package

540   VAST used here has been simulation-tested elsewhere (Grüss *et al.*, 2019; Johnson *et al.*, 2019;

541   Thorson *et al.*, 2019; Brodie *et al.*, 2020), but we recommend ongoing, independent testing to

542   understand the potential impacts of model mis-specification. Importantly, new forms of mis-

543   specification are plausible in seasonal models, e.g., varying rates of temporal change and/or

544   spatial clustering in different seasons.

545   Given rapid changes in climate and its effects on the timing of seasonal processes of growth,

546   migration, mortality, and reproduction, approaches for modeling processes that can handle

547   seasonal and spatial processes are increasingly needed.   We foresee a wide range of future

548   applications for spatial models that include both seasonal and interannual variation. The model

549   is implemented generically within publicly available software that also includes capacity for

550   multivariate analysis, climate-linkages, and skillful near-term forecasting (Thorson, 2019b). We

551   hope that future users will contribute to documenting the benefits and drawbacks of this

552   promising class of models for ecosystem science and management.

553   **Acknowledgments**

570

571

Works cited

Anderson, B. J., Akçakaya, H. R., Fordham, D. A., Martinez-Meyer, E., Thuiller, W., and Brook, B. W. 2009. Dynamics of range margins for metapopulations under climate change. Proceedings of the Royal Society B: Biological Sciences, 276: 1415–1420.

Baier, C. T., and Napp, J. M. 2003. Climate-induced variability in Calanus marshallae populations. Journal of Plankton Research, 25: 771–782.

Banerjee, S., Gelfand, A. E., Finley, A. O., and Sang, H. 2008. Gaussian predictive process models for large spatial data sets. Journal of the Royal Statistical Society. Series B, Statistical methodology, 70: 825–848.

Brodie, S. J., Thorson, J. T., Carroll, G., Hazen, E. L., Bograd, S., Haltuch, M. A., Holsman, K. K., *et al.* 2020. Trade-offs in covariate selection for species distribution models: a methodological comparison. Ecography, 43: 11–24.

Brown, Z. W., and Arrigo, K. R. 2013. Sea ice impacts on spring bloom dynamics and net primary production in the Eastern Bering Sea. Journal of Geophysical Research: Oceans, 118: 43–62.

Butcher, G. S., Fuller, M. R., McAllister, L. S., and Geissler, P. H. 1990. An Evaluation of the Christmas Bird Count for Monitoring Population Trends of Selected Species. Wildlife Society Bulletin (1973-2006), 18: 129–134.

Callaghan, C. T., Martin, J. M., Major, R. E., and Kingsford, R. T. 2018. Avian monitoring – comparing structured and unstructured citizen science. Wildlife Research, 45: 176–184.

Campbell, R. G., Ashjian, C. J., Sherr, E. B., Sherr, B. F., Lomas, M. W., Ross, C., Alatalo, P., *et al.* 2016. Mesozooplankton grazing during spring sea-ice conditions in the eastern Bering Sea. Deep Sea Research Part II: Topical Studies in Oceanography, 134: 157–172.

Conn, P. B., Thorson, J. T., and Johnson, D. S. 2017. Confronting preferential sampling when analysing population distributions: diagnosis and model-based triage. Methods in Ecology and Evolution, 8: 1535–1546.

Coyle, K. O., Pinchuk, A. I., Eisner, L. B., and Napp, J. M. 2008. Zooplankton species composition, abundance and biomass on the eastern Bering Sea shelf during summer: The potential role of water-column stability and nutrients in structuring the zooplankton community. Deep Sea Research Part II: Topical Studies in Oceanography, 55: 1775–1791.

Coyle, K. O., and Gibson, G. A. 2017. Calanus on the Bering Sea shelf: probable cause for population declines during warm years. Journal of Plankton Research, 39: 257–270.

Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., and Wikle, C. K. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. Ecological Applications, 19: 553–570.

Cressie, N., and Wikle, C. K. 2011. Statistics for spatio-temporal data. John Wiley & Sons, Hoboken, New Jersey.

Currie, J. C., Thorson, J. T., Sink, K. J., Atkinson, L. J., Fairweather, T. P., and Winker, H. 2019. A novel approach to assess distribution trends from fisheries survey data. Fisheries Research, 214: 98–109.

Eisner, L. B., Napp, J. M., Mier, K. L., Pinchuk, A. I., and Andrews, A. G. 2014. Climate-mediated changes in zooplankton community structure for the eastern Bering Sea. Deep Sea Research Part II: Topical Studies in Oceanography, 109: 157–171.

Eisner, L. B., and Yasumiishi, E. M. 2018. Large copepod abundance (observed and modeled) as an indicator of pollock recruitment to age-3 in the southeastern Bering Sea. *In* Ecosystem Cosiderations: 2018 Status of the Eastern Bering Sea Marine Ecosystem. Ed. by S. Zador and E. C. Siddon. Alaska Fisheries Science Center, Seattle, WA.

618   Eisner, L. B., Pinchuk, A. I., Kimmel, D. G., Mier, K. L., Harpold, C. E., and Siddon, E. C. 2018. Seasonal,
619   interannual, and spatial patterns of community composition over the eastern Bering Sea shelf in
620   cold years. Part I: zooplankton. ICES Journal of Marine Science, 75: 72–86.
621   Espinasse, M., Halsband, C., Varpe, Ø., Gislason, A., Gudmundsson, K., Falk-Petersen, S., and Eiane, K.
622   2018. Interannual phenological variability in two North-East Atlantic populations of Calanus
623   finmarchicus. Marine Biology Research, 14: 752–767.
624   Flynn, L., and Hilborn, R. 2004. Test fishery indices for sockeye salmon (Oncorhynchus nerka) as affected
625   by age composition and environmental variables. Canadian Journal of Fisheries and Aquatic
626   Sciences, 61: 80–92.
627   Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A., *et al.*
628   2012. AD Model Builder: using automatic differentiation for statistical inference of highly
629   parameterized complex nonlinear models. Optimization Methods and Software, 27: 1–17.
630   Grieve, B. D., Hare, J. A., and Saba, V. S. 2017. Projecting the effects of climate change on Calanus
631   finmarchicus distribution within the U.S. Northeast Continental Shelf | Scientific Reports.
632   Scientific Reports, 7: 1–12.
633   Grosslein, M. D. 1969. Groundfish survey methods. Bureau of Commercial Fisheries, Biological
634   Laboratory.
635   Grüss, A., Walter, J. F., Babcock, E. A., Forrestal, F. C., Thorson, J. T., Lauretta, M. V., and Schirripa, M. J.
636   2019. Evaluation of the impacts of different treatments of spatio-temporal variation in catch-
637   per-unit-effort standardization models. Fisheries Research, 213: 75–93.
638   Heintz, R. A., Siddon, E. C., Farley, E. V., and Napp, J. M. 2013. Correlation between recruitment and fall
639   condition of age-0 pollock (Theragra chalcogramma) from the eastern Bering Sea under varying
640   climate conditions. Deep Sea Research Part II: Topical Studies in Oceanography, 94: 150–156.
641   Hunt, G. L., Coyle, K. O., Eisner, L. B., Farley, E. V., Heintz, R. A., Mueter, F., Napp, J. M., *et al.* 2011.
642   Climate impacts on eastern Bering Sea foodwebs: a synthesis of new data and an assessment of
643   the Oscillating Control Hypothesis. ICES Journal of Marine Science, 68: 1230–1243.
644   Johnson, K. F., Thorson, J. T., and Punt, A. E. 2019. Investigating the value of including depth during
645   spatiotemporal index standardization. Fisheries Research, 216: 126–137.
646   Kai, M., Thorson, J. T., Piner, K. R., and Maunder, M. N. 2017. Predicting the spatio-temporal
647   distributions of pelagic sharks in the western and central North Pacific. Fisheries Oceanography,
648   26: 569–582.
649   Kanamori, Y., Takasuka, A., Nishijima, S., and Okamura, H. 2019. Climate change shifts the spawning
650   ground northward and extends the spawning period of chub mackerel in the western North
651   Pacific. Marine Ecology Progress Series, 624: 155–166.
652   Kass, R. E., and Steffey, D. 1989. Approximate bayesian inference in conditionally independent
653   hierarchical models (parametric empirical bayes models). Journal of the American Statistical
654   Association, 84: 717–726.
655   Kimmel, D. G., Eisner, L. B., Wilson, M. T., and Duffy-Anderson, J. T. 2018. Copepod dynamics across
656   warm and cold periods in the eastern Bering Sea: Implications for walleye pollock (Gadus
657   chalcogrammus) and the Oscillating Control Hypothesis. Fisheries Oceanography, 27: 143–158.
658   Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. M. 2016. TMB: Automatic Differentiation
659   and Laplace Approximation. Journal of Statistical Software, 70: 1–21.
660   Lauth, R. R., and Conner, J. 2016. Results of the 2013 eastern Bering Sea continental shelf bottom trawl
661   survey of groundfish and invertebrate resources. NOAA Technical Memorandum, NMFS-AFSC-
662   331. Seattle, WA.
663   Legault, C. M., Alade, L., Gross, W. E., and Stone, H. H. 2014. Stock Assessment of Georges Bank
664   Yellowtail Flounder for 2014. 2014/01. Transboundary Resources Assessment Committee.
665   https://www.nefsc.noaa.gov/saw/trac/TRD_2014_01_E_.pdf.

666 Legault, C. M., and McCurdy, Q. M. 2018. Stock Assessment of Georges Bank Yellowtail Flounder for
667   2018. 2018/03. Transboundary Resources Assessment Committee.
668   https://www.nefsc.noaa.gov/assessments/trac/documents/gbyt-assessment-2018-v3.pdf.
669 Lindgren, F., Rue, H., and Lindström, J. 2011. An explicit link between Gaussian fields and Gaussian
670   Markov random fields: the stochastic partial differential equation approach. Journal of the Royal
671   Statistical Society: Series B (Statistical Methodology), 73: 423–498.
672 Lindgren, F. 2012. Continuous domain spatial models in R-INLA. The ISBA Bulletin, 19: 14–20.
673 Lindgren, F., and Rue, H. 2015. Bayesian spatial modelling with r-inla. Journal of Statistical Software, 63:
674   1–25.
675 Livingston, P. A., Aydin, K., Buckley, T. W., Lang, G. M., Yang, M.-S., and Miller, B. S. 2017. Quantifying
676   food web interactions in the North Pacific – a data-based approach. Environmental Biology of
677   Fishes, 100: 443–470.
678 Miller, T. J., Das, C., Politis, P., Miller, A., Lucey, S., Legault, C., Brown, R., *et al.* 2010. Estimation of
679   albatross iv to henry b. Bigelow calibration factors. Northeast Fisheries Science Center
680   Reference Document 10-05. Northeast Fisheries Science Center, Woods Hole, MA.
681 Murawski, S. A., and Gavaris, S. 2004. Computation of Allocation Shares for Canada and the USA of the
682   Transboundary Resources of Atlantic Cod, Haddock and Yellowtail Founder on Georges Bank.
683   Transboundary Resources Assessment Committee.
684 Nichol, D. G., Kotwicki, S., Wilderbuer, T. K., Lauth, R. R., and Ianelli, J. N. 2019. Availability of yellowfin
685   sole Limanda aspera to the eastern Bering Sea trawl survey and its effect on estimates of survey
686   biomass. Fisheries Research, 211: 319–330.
687 O'Boyle, R., Beanlands, D., Fanning, P., Hunt, J., Hurley, P., Lambert, T., Simon, J., *et al.* 1995. An
688   overview of joint science/industry surveys on the Scotian Shelf, Bay of Fundy, and Georges Bank.
689   Canadian Technical Report of Fisheries and Aquatic Sciences, 95/133. Department of Fisheries
690   and Oceans. http://publications.gc.ca/collections/collection_2015/mpo-dfo/Fs97-6-3083-
691   eng.pdf.
692 Ono, K., Ianelli, J. N., McGilliard, C. R., and Punt, A. E. 2018. Integrating data from multiple surveys and
693   accounting for spatio-temporal correlation to index the abundance of juvenile Pacific halibut in
694   Alaska. ICES Journal of Marine Science, 75: 572–584.
695 Pedersen, M. W., and Berg, C. W. 2017. A stochastic surplus production model in continuous time. Fish
696   and Fisheries, 18: 226–243.
697 Persson, J., Stige, L. C., Stenseth, N. C., Usov, N., and Martynova, D. 2012. Scale-dependent effects of
698   climate on two copepod species, Calanus glacialis and Pseudocalanus minutus, in an Arctic-
699   boreal sea. Marine Ecology Progress Series, 468: 71–83.
700 Pinsky, M. L., Worm, B., Fogarty, M. J., Sarmiento, J. L., and Levin, S. A. 2013. Marine taxa track local
701   climate velocities. Science, 341: 1239–1242.
702 Pinsky, M. L., and Palumbi, S. R. 2013. Meta-analysis reveals lower genetic diversity in overfished
703   populations. Molecular Ecology. http://onlinelibrary.wiley.com/doi/10.1111/mec.12509/full
704   (Accessed 19 December 2013).
705 Pinto, C., Travers-Trolet, M., Macdonald, J. I., Rivot, E., and Vermard, Y. 2018. Combining multiple data
706   sets to unravel the spatiotemporal dynamics of a data-limited fish stock. Canadian Journal of
707   Fisheries and Aquatic Sciences, 76: 1338–1349.
708 Politis, P. J., Galbraith, J. K., Kostovick, P., and Brown, R. W. 2014. Northeast Fisheries Science Center
709   bottom trawl survey protocols for the NOAA Ship Henry B. Bigelow.
710   https://repository.library.noaa.gov/view/noaa/4825 (Accessed 4 September 2019).
711 R Core Team. 2017. R: A Language and Environment for Statistical Computing. R Foundation for
712   Statistical Computing, Vienna, Austria. https://www.R-project.org/.

713   Rogers, L. A., Griffin, R., Young, T., Fuller, E., Martin, K. St., and Pinsky, M. L. 2019. Shifting habitats
714   expose fishing communities to risk under climate change | Nature Climate Change. Nature
715   Climate Change, 9: 512–516.
716   Rogers, L. A., and Dougherty, A. B. 2019. Effects of climate and demography on reproductive phenology
717   of a harvested marine fish population. Global Change Biology, 25: 708–720.
718   Scranton, K., and Amarasekare, P. 2017. Predicting phenological shifts in a changing climate.
719   Proceedings of the National Academy of Sciences, 114: 13212–13217.
720   Selden, R. L., Thorson, J. T., Samhouri, J. F., Bograd, S. J., Brodie, S., Carroll, G., Haltuch, M. A., *et al.*
721   2020. Coupled changes in biomass and distribution drive trends in availability of fish stocks to US
722   West Coast ports. ICES Journal of Marine Science, 77: 188–199.
723   Siddon, E. C., Kristiansen, T., Mueter, F. J., Holsman, K. K., Heintz, R. A., and Farley, E. V. 2013. Spatial
724   Match-Mismatch between Juvenile Fish and Prey Provides a Mechanism for Recruitment
725   Variability across Contrasting Climate Conditions in the Eastern Bering Sea. PLOS ONE, 8:
726   e84526.
727   Sigler, M. F., Renner, M., Danielson, S. L., Eisner, L. B., Lauth, R. R., Kuletz, K. J., Logerwell, E. A., *et al.*
728   2011. Fluxes, Fins, and Feathers: Relationships Among the Bering, Chukchi, and Beaufort Seas in
729   a Time of Climate Change. Oceanography, 24: 250–265.
730   Sigler, M. F., Stabeno, P. J., Eisner, L. B., Napp, J. M., and Mueter, F. J. 2014. Spring and fall
731   phytoplankton blooms in a productive subarctic ecosystem, the eastern Bering Sea, during
732   1995–2011. Deep Sea Research Part II: Topical Studies in Oceanography, 109: 71–83.
733   Sigler, M. F., Napp, J. M., Stabeno, P. J., Heintz, R. A., Lomas, M. W., and Hunt, G. L. 2016. Variation in
734   annual production of copepods, euphausiids, and juvenile walleye pollock in the southeastern
735   Bering Sea. Deep Sea Research Part II: Topical Studies in Oceanography, 134: 223–234.
736   Skaug, H., and Fournier, D. 2006. Automatic approximation of the marginal likelihood in non-Gaussian
737   hierarchical models. Computational Statistics & Data Analysis, 51: 699–709.
738   Stone, H. H., Gavaris, S., Legault, C. M., Neilson, J. D., and Cadrin, S. X. 2004. Collapse and recovery of the
739   yellowtail flounder (Limanda ferruginea) fishery on Georges Bank. Journal of Sea Research, 51:
740   261–270.
741   Stone, H. H., and Gross, W. E. 2012. Review of the Georges Bank Research Vessel Survey Program, 1987
742   2011. Canadian Manuscript Report of Fisheries and Aquatic Sciences, 2988.
743   Strasburger, W. W., Hillgruber, N., Pinchuk, A. I., and Mueter, F. J. 2014. Feeding ecology of age-0
744   walleye pollock (Gadus chalcogrammus) and Pacific cod (Gadus macrocephalus) in the
745   southeastern Bering Sea. Deep Sea Research Part II: Topical Studies in Oceanography, 109: 172–
746   180.
747   Thorson, J. T., Shelton, A. O., Ward, E. J., and Skaug, H. J. 2015. Geostatistical delta-generalized linear
748   mixed models improve precision for estimated abundance indices for West Coast groundfishes.
749   ICES Journal of Marine Science: Journal du Conseil, 72: 1297–1310.
750   Thorson, J. T., Pinsky, M. L., and Ward, E. J. 2016a. Model-based inference for estimating shifts in species
751   distribution, area occupied and centre of gravity. Methods in Ecology and Evolution, 7: 990–
752   1002.
753   Thorson, J. T., Ianelli, J. N., Larsen, E. A., Ries, L., Scheuerell, M. D., Szuwalski, C., and Zipkin, E. F. 2016b.
754   Joint dynamic species distribution models: a tool for community ordination and spatio-temporal
755   monitoring. Global Ecology and Biogeography, 25: 1144–1158.
756   Thorson, J. T., Fonner, R., Haltuch, M. A., Ono, K., and Winker, H. 2017. Accounting for spatiotemporal
757   variation and fisher targeting when estimating abundance from multispecies fishery data.
758   Canadian Journal of Fisheries and Aquatic Sciences, 74: 1794–1807.

759 Thorson, J. T. 2018. Three problems with the conventional delta-model for biomass sampling data, and a
760   computationally efficient alternative. Canadian Journal of Fisheries and Aquatic Sciences, 75:
761   1369–1382.
762 Thorson, J. T. 2019a. Measuring the impact of oceanographic indices on species distribution shifts: The
763   spatially varying effect of cold-pool extent in the eastern Bering Sea. Limnology and
764   Oceanography, 64: 2632–2645.
765 Thorson, J. T., Adams, G., and Holsman, K. 2019. Spatio-temporal models of intermediate complexity for
766   ecosystem assessments: A new tool for spatial fisheries management. Fish and Fisheries, 20:
767   1083–1099.
768 Thorson, J. T. 2019b. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST)
769   package in stock, ecosystem, habitat and climate assessments. Fisheries Research, 210: 143–
770   161.
771 Usov, N., Kutcheva, I., Primakov, I., and Martynova, D. 2013. Every species is good in its season: Do the
772   shifts in the annual temperature dynamics affect the phenology of the zooplankton species in
773   the White Sea? Hydrobiologia, 706: 11–33.
774 Wilson, R. J., Banas, N. S., Heath, M. R., and Speirs, D. C. 2016. Projected impacts of 21st century climate
775   change on diapause in Calanus finmarchicus. Global Change Biology, 22: 3332–3340.
776 Zipkin, E. F., Ries, L., Reeves, R., Regetz, J., and Oberhauser, K. S. 2012. Tracking climate impacts on the
777   migratory monarch butterfly. Global Change Biology, 18: 3039–3049.
778

779 Table 1 – List of estimated parameters (and associated symbols) governing spatial and temporal

780 variance (listing estimate "Est." and standard error "SE") for each case-study application; this

781 corresponds to all estimated fixed effects except intercepts μ_β, $\beta_{nu}(u)$, and $\beta_{ny}(y)$. Parameters

782 listed as "-" correspond to terms that are dropped due to the corresponding variance approaching

783 zero (see Table S3 for details). Note that the eastern Bering Sea application does not include

784 spatial or temporal variation in the first linear predictor, due to a near 100% encounter

785 probability; linear transformation **H** governing geometric anisotropy involves estimating two

786 parameters which are listed first for each case-study.

Parameter	Symbol	Northwest Atlantic		Eastern Bering Sea	
		Est.	SE	Est	SE
Parameter #1 in **H**	-	0.264	0.083	0.421	0.082
Parameter #2 in **H**	-	0.571	0.097	-0.221	0.088
SD for spatial variation in n	$\sigma_{n\omega}$	1.76	0.182	-	-
SD for spatial season-year interaction in n	$\sigma_{n\varepsilon}$	0.47	0.05	-	-
Natural log. for decorrelation rate in n	$\ln(\kappa_n)$	-3.296	0.096		-
SD for intercept season-year interaction in n	$\sigma_{n\beta}$	-	-	-	-
Autocorrelation for intercepts in n	$\rho_{n\beta}$	-	-	-	-
Autocorrelation for spatial season-year interaction in n	$\rho_{n\varepsilon}$	0.874	0.031	-	-
Natural log. for SD in spatial season and year effects in n	$\ln(\sigma_{n\xi})$	-0.638	0.135	-	-
SD for spatial variation in w	$\sigma_{w\omega}$	0.479	0.069	1.481	0.170
SD for spatial season-year interaction in w	$\sigma_{w\varepsilon}$	0.89	0.063	1.039	0.055
Natural log. for decorrelation rate in w	$\ln(\kappa_w)$	-2.567	0.116	-4.488	0.063
SD for intercept season-year interaction in w	$\sigma_{w\beta}$	-	-	-	-
Autocorrelation for intercepts in w	$\rho_{w\beta}$	-	-	-	-
Autocorrelation for spatial season-year interaction in w	$\rho_{w\varepsilon}$	-0.149	0.085	0.721	0.034
Natural log. for SD in spatial season and year effects in w	$\ln(\sigma_{w\xi})$	-	-	-0.644	0.163
Natural log. for SD of measurement error	$\ln(\sigma_m)$	-0.133	0.04	-0.103	0.017

787

788

789　　Fig. 1 – Data availability (y-axis) for season (colored line) and each year (x-axis) for case-studies

790　　in the Northwest Atlantic Ocean (top panel) and eastern Bering Sea (bottom panel), where the

791　　colorbar for each season is indicated in each legend.
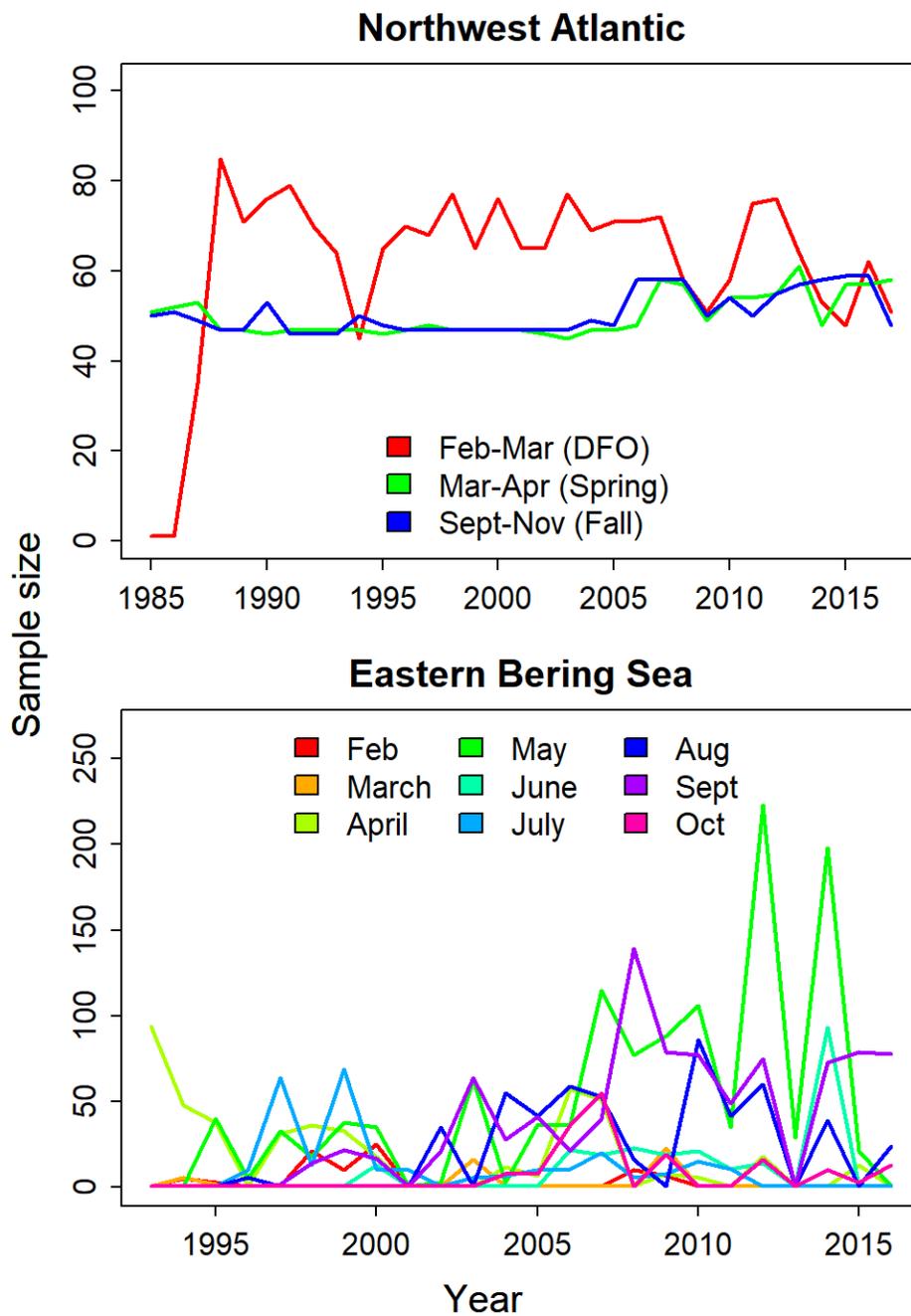
792

793

794   Fig. 2 – Index of relative abundance (y-axis using log-scale; scaled to have mean of 1 across

795   years) for a design-based index (black line), seasonal spatio-temporal model (blue line), or

796   independent spatio-temporal model (red line) including +/- 1 standard error (shaded interval) of

797   yellowtail flounder for 1985-2017 (x-axis) for each of three seasons (DFO, as well as spring and

798   fall NEFSC bottom trawl surveys). Note that the design-based and independent spatio-temporal

799   models cannot estimate abundance for the DFO survey in 1985/1986, and therefore have no

800   estimates for those years (top panel)
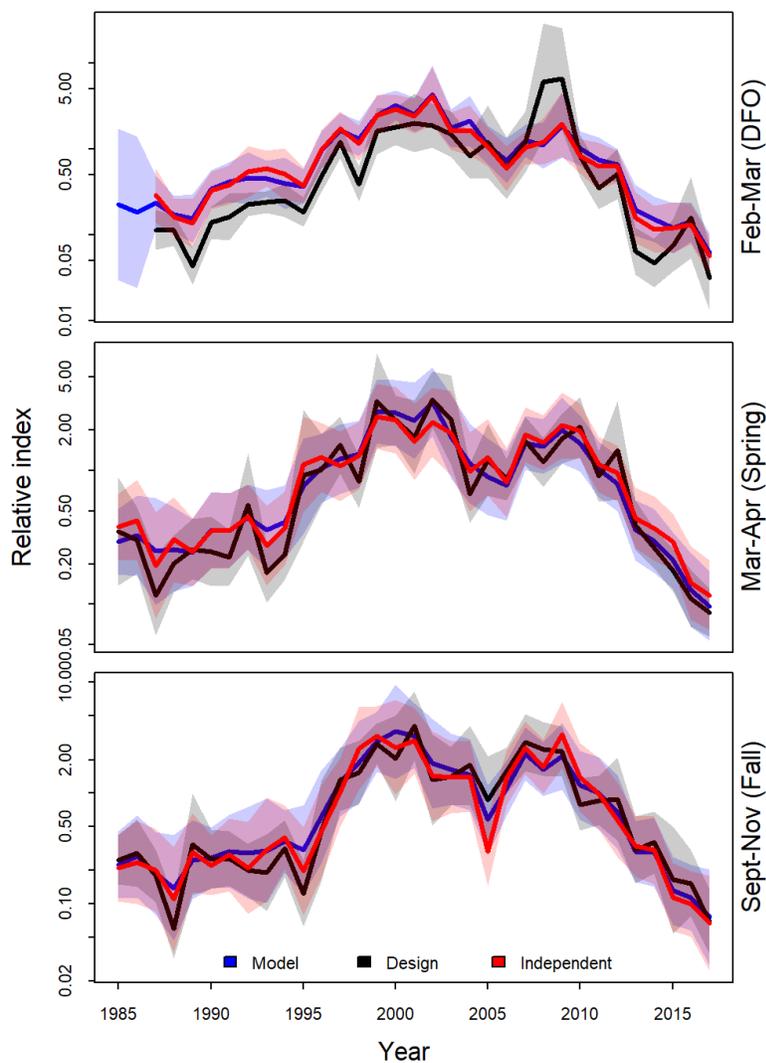
801

802   Fig. 3 – Spatio-temporal variation in the log-linked linear predictor for numbers-density $\varepsilon_n^*(s,t)$

803   for each location s and year-season interval t for yellowtail flounder on Georges Bank in the

804   Northwest Atlantic Ocean for 1989-1992 (rows) of the 24 modeled years (1985-2017) and three

805   seasonal surveys (columns) in each year (DFO Feb.-March, NEFSC Spring March-April, and

806   NEFSC Fall Sept.-Nov. surveys). Each panel also shows the location of available data in that

807   season and year (black dots), which highlights the absence of data from DFO in the southwestern

808   portion of the stock area. Years are selected to highlight the estimates of increased density in the

809   DFO survey in the southwestern portion 1990-1991 which is not sampled by that survey; this

810   estimated hotspot is informed by sharing information among surveys. The use of a log-linked

811   linear predictor means, for example, that a location s and year-season t with a value of

812     $\varepsilon_n^*(s,t) = 0.1$ has an approximately $\exp(0.1) = 10.5\%$ higher prediction of numbers-density

813     $n^*(s,t)$, while a location and time where $\varepsilon_n^*(s,t) = 1.0$ has an approximately $\exp(1.0) = 172\%$

814   higher prediction of numbers-density $n^*(s,t)$ than the value otherwise expected for that location

815   and time.
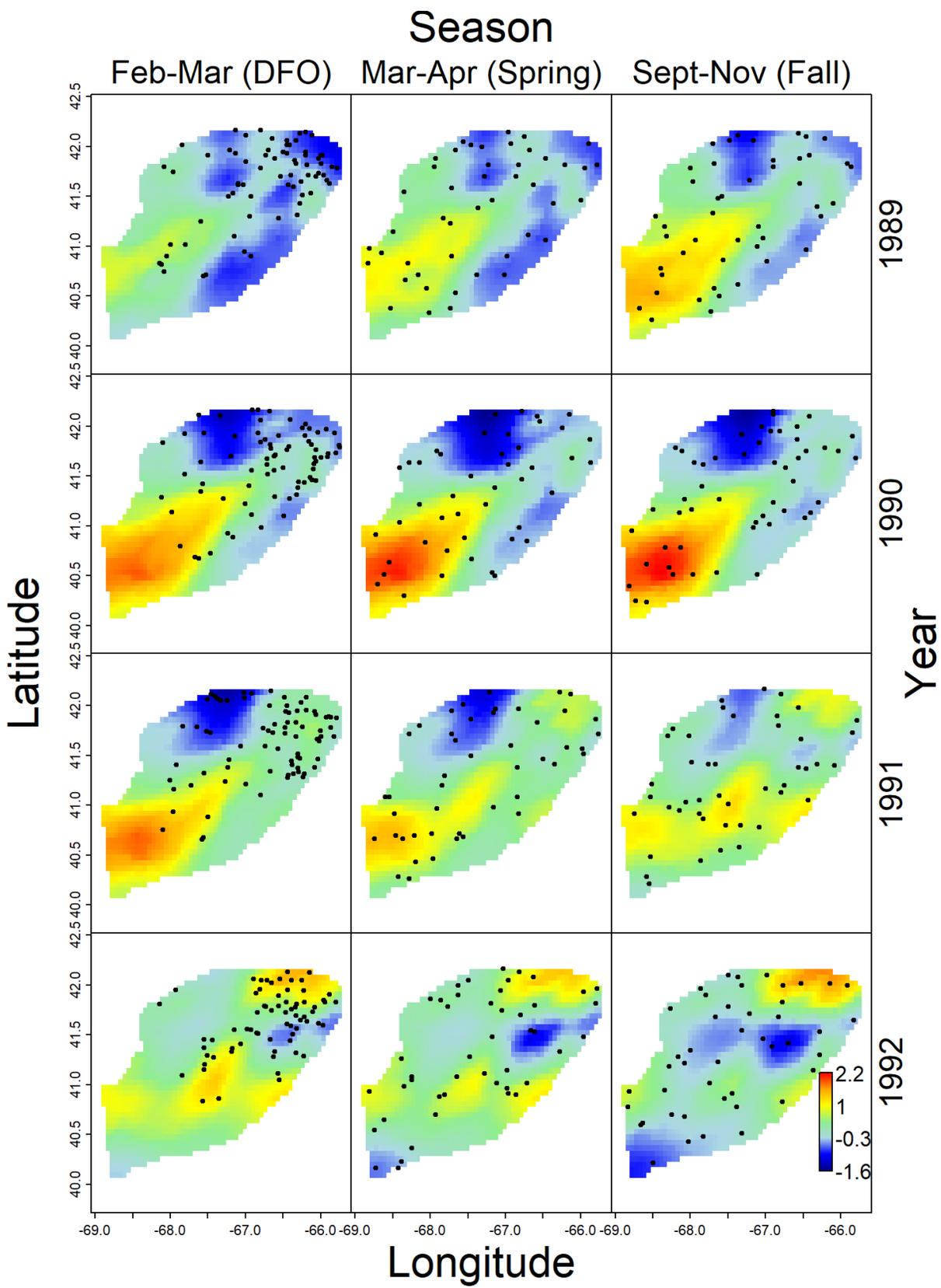
819   Fig. 4 – Natural logarithm of estimated density (vertically-integrated from 5 meters above seafloor to surface), $\ln(number \cdot km^{-2})$,

820   for copepodite stages 3+ *Calanus marshallae/glacialis* for four selected years (rows) from 24 analyzed years (1993-2016) and nine

821   month-long intervals in each year (columns). Each panel also shows the location of available data in that season and year (black dots).
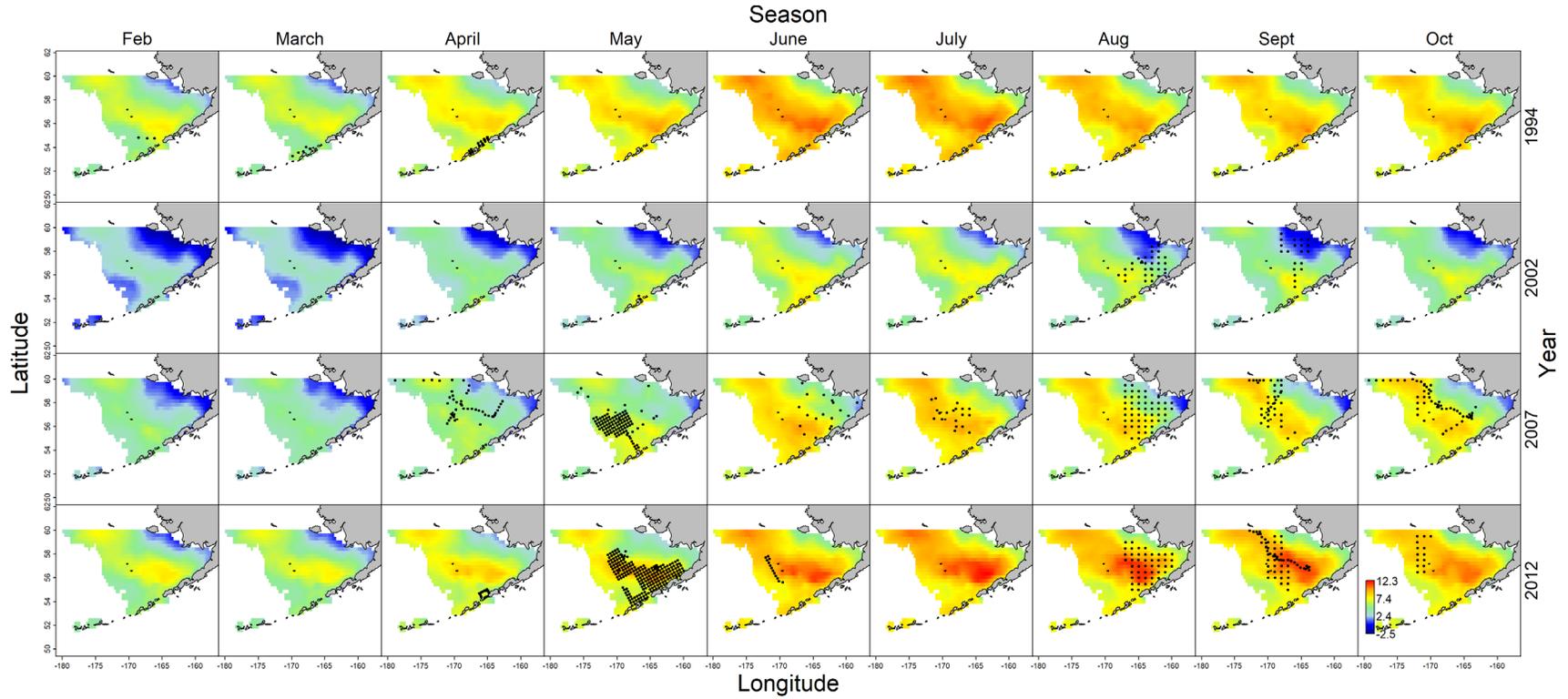
822

823

824   Fig. 5 – Natural logarithm of estimated total abundance (left-hand y-axis) for copepodite stage

825   3+ *Calanus marshallae/glacialis* in Feb.-Oct. (x-axis) for each year (panels) from 1993-2016,

826   estimated using a seasonal spatio-temporal model. Each panel shows an estimate (black line) +/-

827   one standard error (grey shaded interval), and also shows the sample size for each season (dashed

828   line; using right-hand y-axis). Note that the left and right-hand y-axes have identical range for

829   all panels, but abundance (left-hand y-axis) is in log-space while sample size (right-hand y-axis)

830   is in natural space.
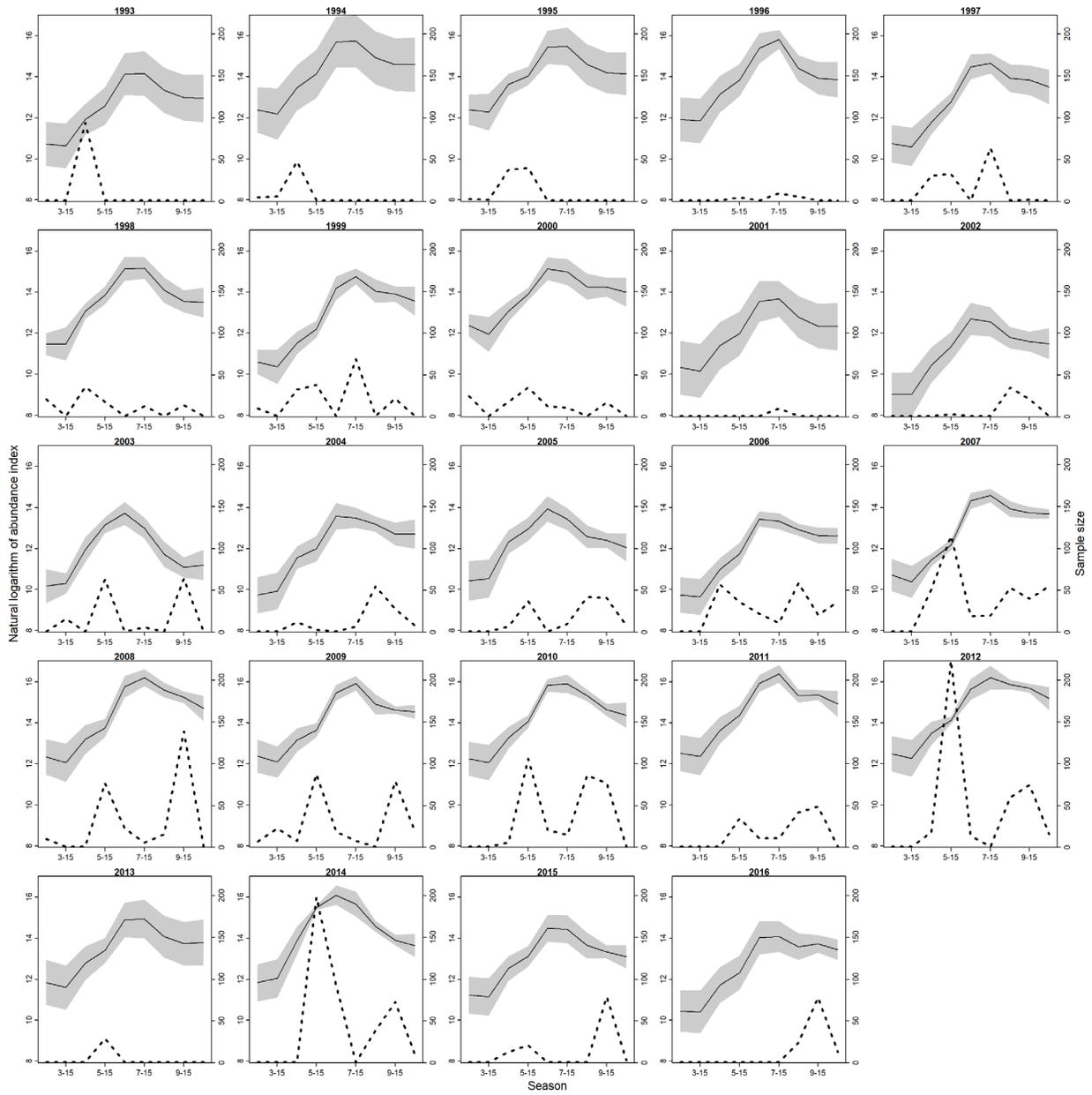
831

832

833   Fig. 6 – Estimate of abundance-weighted average season (calculated using Eq. 11; black line

834   with +/- one standard error as grey shaded area; scale on left-hand y-axis) in every year 1993-

835   2016 based on density estimates in every month Feb.-Oct. (x-axis) from a seasonal spatio-

836   temporal model compared with the summertime spatial extent of cold ($\leq 2°\,C$) near-bottom

837   waters in the eastern Bering Sea (dashed line, "cold pool extent"; scale on right-hand y-axis).

838   We also show the Pearson correlation between cold-pool extent and the abundance-weighted

839   average season, as well as the standard error for this correlation based on 100 correlation values

840   calculated when simulating densities from the joint precision of fixed and random effects

841   (bottom right).

842

843