

A Gaussian Field Approach to Generating Spatial Age Length Keys

Jonathan Babyn^{*a}, Divya Varkey^{†b}, Paul Regular^{‡b}, Danny Ings^{§b}, and Joanna Mills Fleming^{¶a}

^a*Dalhousie University, Department of Mathematics & Statistics, 6316 Coburg Road, PO Box 15000, Halifax NS B3H 4R2*

^b*Fisheries and Oceans Canada, Northwest Atlantic Fisheries Center, St. John's, Newfoundland and Labrador*

2021-03-16

Abstract

Estimating the age composition of a fish population is a critical first step in all stock assessments that apply age-structured models. Often this is done through the use of an Age Length Key (ALK), which links a subsample of fish that have had their ages determined to those that have only had their lengths measured in order to obtain an estimate of the age structure of the entire sample. ALKs can suffer from data gaps and sampling artifacts and are limited in both how they can reflect spatial variability and how spatial information can be incorporated.

We propose a novel spatial ALK model that uses an approximation of a Gaussian Field and has the ability to account for physical barriers (e.g. islands, coastlines) in the study area. Our approach is compared with a previously suggested spatial ALK model as well as non-spatial approaches using both real and simulated survey data. We find that spatial ALK approaches reduce errors in stratified estimates of abundance at age over non-spatial approaches and that incorporating physical barriers can deliver more realistic results.

1 Introduction

Stock assessments are tools that can allow us to understand the overall health of a fish stock. They enable quantifying the abundance, age and length compositions of the population and determine indications of whether the stock is facing overexploitation (Worm et al., 2009). They play a key part in helping to rebuild and maintain fisheries around the world (Worm et al., 2009; Hilborn, Ovando, 2014). Stock assessment methods have evolved from simple methods based only on catch data to models that integrate additional sources of data, to modern state-space approaches (Aeberhard et al., 2018) that allow for increasing levels of inference and precision.

Age structured methods can greatly simplify stock assessment models as ages link directly to the numbers of survivors in each year. However, for most species accurately and easily determining the age of a fish can be a time consuming and expensive process that often requires an expert

*Corresponding Author: jn805248@dal.ca

†Divya.Varkey@dfo-mpo.gc.ca

‡Paul.Regular@dfo-mpo.gc.ca

§Danny.Ings@dfo-mpo.gc.ca

¶Joanna.Flemming@dal.ca

33 counting the number of growth rings on an otolith or similar procedure. Measuring the length of a
 34 fish is much easier, less lethal and it can be done on site for low cost . In order to take advantage of
 35 the benefits of age structured methods, approaches for estimating the age of a fish from its cheaply
 36 measured length are commonly used, like Age Length Keys (ALKs) (Aanes, Vølstad, 2015).

37 ALKs have been used to estimate the age of fish for over 80 years (Fridriksson, 1934). They are
 38 based on the idea that the proportion of fish at age a , p_a is equal to

$$p_a = \sum_{i=1}^K k_i p_{a|i} \quad (1)$$

39 where i indexes discrete length bins $i = 1$ to K and k_i is the proportion of fish in length bin i , $p_{a|i}$
 40 is the observed conditional probability (or proportion) of being age a given membership in length
 41 bin i . An ALK is then simply a matrix of proportions of fish at age a given length i . To convert the
 42 sampled length frequencies to estimates at age, the length frequencies are multiplied by the ALK to
 43 get the numbers at age. An example of a traditional ALK along with an example of a smooth model
 44 based ALK is shown in Table 1. ALKs are often constructed separately for different covariates such
 45 as sex, time of year and gear type depending on the species and application. Traditional ALKs
 46 can easily suffer from data gaps resulting in some fish not being assigned an age estimate. ALKs
 47 also suffer from low sample numbers particularly for rarer older age classes. An example of such
 48 a sampling artifact can be seen in Table 1a where any fish assigned to the 49 cm length bin will
 49 automatically be assumed to be age 9 despite the existence of shorter older fish.

50 Smooth ALKs have non-zero proportions at every possible length bin ensuring that all fish are
 51 assigned an age estimate. Kvist et al., 2000 suggested the possibility of generating smooth ALKs
 52 using a Continuation Ratio Logit (CRL) model. Smooth ALKs have been applied before to lesser
 53 sandeel, (Rindorf, Lewy, 2001) North Sea haddock, (Stari et al., 2010; Berg, Kristensen, 2012) cod,
 54 herring and whiting (Berg, Kristensen, 2012) . They can also mitigate the effects of the sampling
 55 artifacts mentioned above, demonstrated in Table 1b where the probability of being a particular
 56 age given a length of 49cm is more suitably spread across nearby age classes and not just associated
 57 with age nine.

58 Many fish species are known to aggregate in different locations and may be in different spots
 59 at different stages of their lifecycles (Parrish, 1999). They may also have differing levels of growth
 60 depending on location (Punt et al., 2015). Incorporating spatial information into traditional ALKs
 61 requires dividing the study area into subareas. The greater the number of subareas the more
 62 sparse the data, the more likely gaps and other issues are to arise in the ALK resulting in more
 63 missed or incorrectly aged fish. Berg, Kristensen, 2012 presented a way of constructing ALKs
 64 for point referenced data using a Generalized Additive Model (GAM) and thin-plate regression
 65 splines. They found better internal and external consistencies for age based survey indices when
 66 using spatial ALKs, in addition to observing differences in ALKs constructed in different areas.

67 However, the method of Berg, Kristensen, 2012 does not account for boundaries posed by phys-
 68 ical barriers such as landmasses that may be present in the study area. These oversimplifications
 69 results in predictions that smooth under landmasses. That is, the spatial part of the model will
 70 ignore any landmasses and predicted probabilities will ignore marine distance which may result in
 71 poor estimates for samples on opposite sides of a large landmass for instance. This could be a
 72 problem if for example a large group of young fish inhabits a bay as smoothing under landmasses
 73 may artificially increase the probability of young fish living on the other side of the bay.

74 The spatial ALK approach presented here addresses the problem of smoothing over landmasses
 75 by using an approximation of a Gaussian Random Field (GF) that has support for physical barriers
 76 (Bakka, Vanhatalo, J. B. Illian, et al., 2019) . Desirably it still allows ALKs to be constructed at

77 any location within the study area. Approximations of GFs have previously been proposed to help
 78 model spatial indices of abundance (Thorson, Shelton, et al., 2015; Thorson, Barnett, 2017), species
 79 distribution (Bakka, Vanhatalo, J. Illian, et al., 2016) along with other non-marine uses such as
 80 global temperature data (Lindgren, Rue, 2011).

81 [Table 1 about here.]

82 In Section 2 our spatial ALK is fully described and all necessary background knowledge pro-
 83 vided. In Section 3.1 the proposed Gaussian field model with barrier support (GFB) is tested using
 84 simulated survey data to determine the benefit of using a spatial ALK instead of one that ignores
 85 all spatial structure. It is tested alongside a traditional ALK, a non-spatial CRL model and a
 86 spatial GAM ALK implementation. Finally in Section 3.2, the four methods are applied to two
 87 real datasets (Cod and American Plaice) from Fisheries and Oceans Canada (DFO)’s multi-species
 88 bottom trawl Research Vessel (RV) survey. The method proposed is implemented as an R package
 89 called `barrierALK` and is available on Github (<https://github.com/jgbabyn/barrierALK>).

90 2 Methods

91 2.1 Ordinal Regression and Continuation Ratio Logits

92 An ALK can be constructed using any classifier capable of handling multiple age classes and
 93 returning probabilities of a fish with a particular set of covariates (e.g. male, caught using gillnet,
 94 etc.) belonging to each age class. The ALK is simply those probabilities for every set of observed
 95 covariates. Ages naturally have an ordering associated with them, a seven year old fish must have
 96 first been a six year old fish and before that a five year old fish and so on. Ordinal regression can
 97 handle ordered categorical data and return class probabilities (Agresti, 2003). It is also possible
 98 to incorporate spatial structure directly into ordinal regression models through the use of splines
 99 or random fields. This is not the case for some alternative multi-class classifiers like Classification
 100 and Regression Trees (CARTs). A number of different ordinal regression methods exist such as
 101 cumulative logits, adjacent-category logits, etc. Continuation Ratio Logits (CRLs) are one method
 102 that offers a few advantages (Agresti, 2003)(Harrell, 2014, p. 311-312, 321–322).

103 CRLs models have the advantage over other ordinal regression methods in that it is very easy
 104 to remove or loosen the proportional odds assumption. This allows for all covariates or some subset
 105 of covariates to be able to freely vary with every level of a category (Harrell, 2014, p. 319-322).
 106 In addition CRL models can be represented using $\mathcal{A} - 1$ binomial models which allows their fitting
 107 using any software capable of performing logistic regression(Agresti, 2003).

108 Suppose the i th aged fish is observed to be age a , and there are \mathcal{A} total ages. The CRL for the
 109 i th observation with the corresponding vector of covariates, \mathbf{x}_i , that must of course include length
 110 along with others (e.g. sex, gear type, etc.), is

$$\text{logit}(\pi_a[\mathbf{x}_i]) = P(X_i = a | X_i \geq a) \tag{2}$$

111 where

$$\pi_a[\mathbf{x}_i] = \frac{p_a[\mathbf{x}_i]}{p_a[\mathbf{x}_i] + \dots + p_{\mathcal{A}}[\mathbf{x}_i]} \tag{3}$$

112 and p_a is the proportion of fish at age a . In other words, the CRL for the i th observation is the
 113 probability of being age a given it is at least age a or greater (Agresti, 2003).

114 The unconditional probabilities $P(X_i = a)$ can be found by

$$\begin{cases} \pi_a[\mathbf{x}_i], & a = R \\ \pi_a[\mathbf{x}_i] \sum_R^{a-1} (1 - \pi_i[\mathbf{x}_i]) & R < a < A \\ 1 - \sum_R^{A-1} (1 - \pi_i[\mathbf{x}_i]) & a = A \end{cases} \quad (4)$$

115 where R is the first age in the model, and A is the last. In aging data R can refer to the age
 116 of recruitment to the survey or fishery, A identifies a plus group or the final age involved (Berg,
 117 Kristensen, 2012). In the models presented here, length has an individual parameter for each
 118 age group which allows for greater flexibility. This is known as relaxing the proportional odds
 119 assumption.

120 2.2 Random Fields

121 Random fields are collections of random variables, $\{X(\mathbf{s}), \mathbf{s} \in \mathbf{S}\}$. \mathbf{S} is the set of indices and \mathbf{s} is
 122 the index (Ross, 2014). Typically a set of indices is a set of locations, but could also incorporate
 123 time. The index set \mathbf{S} can be a discrete set, continuous, finite or infinite. The random variables
 124 in random fields can follow any of the typical distributions used such as Gaussian, Student's t ,
 125 Gamma, etc. Gaussian Random Fields (GFs) are those in which all the $X(\mathbf{s})$ are normally or
 126 Gaussian distributed (Rue, Held, 2005). A GF can be specified by its mean function $\mu(s)$ and
 127 covariance function $\text{Cov}(s, t)$, $s, t \in \mathbf{S}$. A popular choice of covariance function for spatial data
 128 is the Matérn covariance function,

$$c(s, t) = \sigma_u^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{8\nu} \frac{\|s - t\|}{r} \right) K_\nu \left(\sqrt{8\nu} \frac{\|s - t\|}{r} \right) \quad (5)$$

129 where Γ is the Gamma function, ν is a smoothness parameter, K_ν is the modified Bessel function
 130 of the second kind with order ν , r is the range parameter which is the spatial distance when the
 131 correlation is ≈ 0.13 , σ_u is the marginal standard deviation (Bakka, Rue, et al., 2018) and $\|s - t\|$
 132 is the distance between two points.

133 Using a GF directly is not computationally tractable for large problems. The cost to factorize the
 134 resulting dense covariance matrix is cubic in time. As a result a number of alternative approaches
 135 to try and get around the high computational cost have been proposed. Gaussian Markov Random
 136 Fields (GMRFs) are GFs with the Markov property, that is

$$P(\mathbf{s}_i | \{\mathbf{s}_j : j \neq i\}) = P(\mathbf{s}_i | \mathbf{x}_j : j \in \mathcal{N}_i) \quad (6)$$

137 where the neighbours \mathcal{N}_i to the location \mathbf{s}_i are the points $\{\mathbf{s}_j, j \in \mathcal{N}_i\}$ that are close or connected to
 138 \mathbf{s}_i . Conditional on its neighbours the mean at a location is independent of all other locations. The
 139 Markov property ensures that the precision matrix (inverse of the covariance matrix) is sparse. The
 140 sparsity of the precision matrix reduces the memory needed and overall computational time required
 141 (Rue, Held, 2005). However GMRFs have traditionally been limited for spatial applications as they
 142 require that areas be broken into predefined regions beforehand. This may be difficult to do in
 143 practice and it also limits the spatial resolution available.

144 2.2.1 Gaussian Random Field Approximation using Stochastic Partial Differential 145 Equations

146 Lindgren et al. (2011) found an explicit link between GFs and GMRFs when using a Matérn
 147 covariance function. A valid positive semi-definite covariance matrix is the result of the solution

148 of a set of Stochastic Partial Differential Equations (SPDEs), which creates an approximation of a
 149 GF using a GMRF. This allows the benefits of modelling as a GF with the computational speed of
 150 a GMRF.

151 The SPDE method requires creating a Delaunay triangulation or mesh of the study area such
 152 as through R-INLA’s `inla.mesh.2d` function, which can then be used to generate the matrices that
 153 are used for the Finite Element Method (FEM) solution to the SPDE. The Matérn field is the
 154 solution $u(\mathbf{s})$ to the SPDE

$$u(\mathbf{s}) - \nabla \cdot \frac{r^2}{8} \nabla u(\mathbf{s}) = r \sqrt{\frac{\pi}{2}} \sigma_u \mathcal{W}(\mathbf{s}) \quad (7)$$

155 assuming that the smoothness parameter $\nu = 1$, where ∇ is defined as $\left(\frac{\delta}{\delta x}\right)$, r is the range
 156 parameter, σ_u is the marginal standard deviation of the model component u , $\mathcal{W}(\mathbf{s})$ refers to white
 157 noise. The approximation to the spatial GF $\tilde{u}(\mathbf{s})$ is distributed $\mathcal{N}(\mathbf{0}, \mathbf{Q}(\sigma_u, r)^{-1})$ with $\mathbf{Q}(\sigma_u, r)$ is
 158 the precision matrix that results from the FEM solution to the SPDE with hyperparameters σ_u
 159 and r .

160 The mesh helps ensure that the resulting precision matrix is sparse as well. Every node in the
 161 mesh is an element in the resulting covariance matrix. The more nodes in the mesh, the denser it
 162 is and the better the approximation to the GF. This comes as a tradeoff, as computational time
 163 increases non-linearly with the addition of more nodes.

164 Recently Bakka et. al (2019) extended the SPDE GF approximation method of Lindgren et.
 165 al (2011) to support physical barriers in a spatial GF such as the problem presented by coastlines.
 166 Their method has several advantages over other proposed methods of incorporating boundary
 167 information into a spatial model like the soap film smoother proposed by Wood et al., 2008. It
 168 is robust to the selection of boundary polygons, takes similar amounts of computational time and
 169 is not particularly hard for the applied practitioner to use beyond defining the mesh and barrier
 170 polygons (Bakka, Vanhatalo, J. B. Illian, et al., 2019). Under the assumption that the smoothness
 171 parameter $\nu = 1$, the barrier Matérn field is the solution $u(\mathbf{s})$ to the SPDE

$$u(\mathbf{s}) - \nabla \cdot \frac{r^2}{8} \nabla u(\mathbf{s}) = r \sqrt{\frac{\pi}{2}} \sigma_u \mathcal{W}(\mathbf{s}) \quad \text{for } s \in \Omega_n \quad (8)$$

172

$$u(\mathbf{s}) - \nabla \cdot \frac{r_b^2}{8} \nabla u(\mathbf{s}) = r_b \sqrt{\frac{\pi}{2}} \sigma_u \mathcal{W}(\mathbf{s}) \quad \text{for } s \in \Omega_b \quad (9)$$

173 where Ω_n is the set of nodes outside the boundary, Ω_b is the set of nodes inside the boundary, and
 174 r_b is not a new range parameter but instead a predetermined fraction of r . In this case $r_b = \frac{1}{10}r$.
 175 This has the effect of essentially making the decorrelation range close to zero for points that fall
 176 within the barrier creating the desired boundary properties. Other parameters are the same as
 177 described above. Further details on solving the SPDE for the GF approximation can be found in
 178 Lindgren, Rue, 2011, Bakka, Vanhatalo, J. B. Illian, et al., 2019 and Bakka, 2018.

179 Prediction and fitting of points that do not fall exactly on mesh node are handled by projector
 180 matrix \mathbf{A} . \mathbf{A} is also a sparse matrix that has the same number of rows as data being predicted
 181 or fit and a column for every node in the mesh. Every row of \mathbf{A} has either one or three non-zero
 182 entries. If the data point falls exactly at a mesh node then the non-zero entry will be 1 at that
 183 node’s column. For points not at a mesh node, the three non-zero entries are the distances from
 184 the three vertices of the triangle in the mesh that the point lies in. The \mathbf{A} matrix is multiplied
 185 against the observed random effects for nodes in the mesh and estimates of the random effects at
 186 each point are found and usable in the model (Bakka, Vanhatalo, J. B. Illian, et al., 2019).

187 **2.3 GF Spatial Age-Length Key**

188 The GFB model presented here and implemented in `barrierALK` combines the CRL and barrier
 189 approach together. This GFB model is

$$\text{logit}(\pi_a[\mathbf{x}_i]) = \alpha_a + \beta_a l_i + \xi_{a,s} \tag{10}$$

190 where α_a is the intercept for age a , β_a is the length parameter for age a and $\xi_{a,s}$ is the spatial
 191 intercept resulting from the GF at location \mathbf{s} :

$$\xi_{a,s} = \begin{cases} \text{MVN}\left(\mathbf{0}, \frac{\sigma_u^2}{(1-\varphi_a^2)}c(s, 0)\right) & a = 1 \\ \text{MVN}\left(\varphi_a \xi_{a-1,s}, \sigma_u^2 c(s, 0)\right) & a > 1. \end{cases} \tag{11}$$

192 The φ_a allows for spatial correlation between age classes, if it exists. What this means in practical
 193 terms is that if age structure varies in space, φ_a can measure how correlated that relationship may
 194 be.

195 **2.4 Estimation**

196 Estimation is performed using the R package Template Model Builder (TMB). TMB uses the
 197 Laplace approximation to approximate the integrals in the log likelihood resulting from the random
 198 effects that need to be integrated out. TMB uses Automatic Differentiation (AD) to generate the
 199 derivatives for a given objective function which can result in a speed up when paired with an
 200 optimizer capable of utilizing derivative information (Kristensen et al., 2016).

201 When fitting ordinal regression models via Maximum Likelihood Estimation (MLE), optimiza-
 202 tion algorithms will often fall into local minima resulting in unrealistic parameter estimates which
 203 also has the effect of reducing the predictive accuracy of the model. Penalizing the log likelihood
 204 can improve parameter estimates for classes with low numbers of observations (Harrell, 2014, p. 323,
 205 209–213). The penalized log likelihood is written as

$$\log L - \frac{1}{2} \lambda \boldsymbol{\beta}' \mathbf{P} \boldsymbol{\beta} \tag{12}$$

206 where L is the likelihood from an unpenalized model, $\boldsymbol{\beta}$ being the vector of the fixed effects co-
 207 efficients, λ the penalty factor chosen by cross validation and \mathbf{P} the penalty matrix. Parameters
 208 relating to continuous variables are scaled in \mathbf{P} by their standard deviation, parameters relating
 209 to categorical variables use the penalty function $\sum_i^c (\beta_{fi} - \bar{\beta}_f)^2$ where f is a categorical variable
 210 in the model with c levels, $\bar{\beta}_f$ is the mean of all c β_{fi} . This shrinks parameters towards the mean
 211 parameter value which avoids biasing towards a specific level (Harrell, 2014, p. 209-213)(Verweij,
 212 Van Houwelingen, 1994).

213 The optimal choice of λ can be chosen by k-fold cross validation (CV). This can be quite time
 214 intensive for large spatial models. Due to the presence of random effects in the spatial ALK model,
 215 an approach like Generalized Cross Validation (GCV) which would avoid the time consuming k-fold
 216 CV can also not be used due to the difficulty in finding an influence matrix if it even exists at all.
 217 Instead a modified Akaike Information Criterion (AIC) is proposed here in place of k-fold CV. The
 218 modified AIC used here is defined to be

$$\text{LR } \chi^2 - \text{effective degrees of freedom} \tag{13}$$

219 where LR χ^2 is the likelihood ratio value comparing the null model containing only an intercept
 220 to the model with the final penalized parameters ignoring the penalty function and the effective
 221 degrees of freedom that result from taking the penalization into account that are found by

$$\text{trace}(I(\boldsymbol{\beta}^P)V(\boldsymbol{\beta}^P)) \tag{14}$$

222 where $\boldsymbol{\beta}^P$ is the vector of penalized fixed effects parameters resulting from MLE, $I(\boldsymbol{\beta}^P)$ is the
 223 information matrix resulting from the model using the penalized parameters but ignoring the
 224 penalty function and $V(\boldsymbol{\beta}^P)$ is the covariance matrix for the penalized parameters of the model
 225 when taking the penalty function into account (Gray, 1992; Verweij, Van Houwelingen, 1994). The
 226 model that maximizes the modified AIC over the selection of possible λ values is the version of the
 227 model most likely to result in the best predictive accuracy for a new data set. This method has been
 228 shown to be asymptotically equivalent to CV approaches for selecting the penalty factor (Harrell,
 229 2014, p. 209-213). While the modified AIC approach does not explicitly account for random
 230 effects in the model, comparisons are only made between models based on the same data and with
 231 the same number of random effects, only the value of λ changes. The values of λ considered for
 232 penalization for the non-spatial CRL and the proposed GFB model are 0, 0.001, 0.01, 0.1, 0.25,
 233 0.5, 1, 2, 5 and 10.

234 2.5 Simulation Study

235 Survey data was simulated using a modified version of the `SimSurvey` R package available on GitHub.
 236 `SimSurvey` was originally designed with the purpose of testing different stratified random sampling
 237 survey designs for a research vessel survey aimed at estimating abundance. `SimSurvey` is capable
 238 of generating data quite similar to those resulting from a stratified random survey design like the
 239 multi-species bottom trawl research vessel survey that DFO performs annually in the Newfoundland
 240 region and elsewhere. With simulated data it's possible to know the true age structure and the
 241 abundance numbers for the population. Further details on the the simulation study are provided
 242 in Appendix A (Regular et al., 2020).

243 The population simulated with `SimSurvey` is similar to the cod population living in Northwest
 244 Atlantic Fisheries Organization (NAFO) subdivision 3Ps. The fish were set to grow according to a
 245 Von Bertalanffy growth curve with an asymptotic length (L_∞) of 120 cm and K parameter of 0.5.
 246 The population is spatially distributed within grid cells grouped into strata based on depth and
 247 a stratified random survey is taken by sampling random cells. A subsample of fish from sampled
 248 locations is obtained based on length stratified sampling and considered to have been "aged". This
 249 subsample of fish is what is used to construct the four different ALKs described below.

250 Spatial methods have an opportunity to improve the estimate of age structure by being better
 251 able to discriminate between age classes with overlapping length distributions by taking into account
 252 the location where sampling occurred. The simulated population does not let length at age vary
 253 from location to location rather the distribution of age classes varies spatially.

254 The four different aging methods applied to the simulated survey data are the traditional ALK,
 255 a smooth ALK made from a CRL model involving only length as a covariate,

$$\text{logit}(\pi_a[\mathbf{x}_i]) = \alpha_a + \beta_a l_i \tag{15}$$

256 a GAM based model similar to the one presented in Berg, Kristensen, 2012,

$$\text{logit}(\pi_a[\mathbf{x}_i]) = \alpha_a + \beta_a l_i + f(\mathbf{s}) \tag{16}$$

257 where f is a function of location \mathbf{s} using thin-plate regression splines and finally the model proposed
 258 here as described in Equation 10. An automatic selection of the maximum basis dimension (k) is

259 used for the thin-plate regression splines with AIC smoothness selection. The automatic selection
 260 is based on the method used by the DATRAS package discussed in Berg, Kristensen, 2012 where
 261 the maximum basis dimension is the number of unique observations of the covariates appearing in
 262 the smooth terms minus one (i.e the number of unique (non-zero) tow locations minus one) unless
 263 there are less than 10 unique locations for which then it falls back to a GLM. If including spatial
 264 information increases the accuracy of predicting what age class a fish belongs to, then the stratified
 265 survey estimates of abundance at age should also be closer to true abundance numbers at age. For
 266 all of the approaches age 10 was taken to be a plus group.

267 The data were simulated in a simplified area with a large peninsula-like landmass represented
 268 by rectangles. A plot of the landmass and an example simulation mesh can be found in Appendix
 269 A. In practice any physical boundary can be defined with the only caveat being that more detailed
 270 boundaries require more nodes in the mesh and increase the computational time required. A sim-
 271 plified boundary was chosen for computational convenience to reduce the required time needed to
 272 run the model hundreds of times. 500 simulated surveys were performed, and stratified survey
 273 estimates of abundance were created for each survey using the same methods outlined in (Smith,
 274 Somerton, 1981) and the same method applied to DFO’s bottom trawl multi-species survey con-
 275 ducted in the Newfoundland region annually (Ings et al., 2019). The simulation does not account
 276 for observation error and each simulation has a new realization of the true abundance in each run.
 277 The Root Mean Squared Error (RMSE) is calculated on the true total population at age available
 278 to the survey (adjusting for selectivity) and the stratified survey estimates resulting from the age
 279 frequencies made from each of the four methods, the GFB, GAM, non-spatial CRL and traditional
 280 ALK.

281 For the simulated survey 96 tows were conducted per survey in 48 strata based on depth. A
 282 mean of 2774.26 simulated fish were caught and measured in each survey with a mean of 454.38 of
 283 those being “aged” and used for constructing the ALKs. Length stratified sampling was used in
 284 selecting the subsamples to be aged. Out of 500 simulations, 73 failed to due to too low sampling
 285 numbers. Specifically in those simulations, zero catches of older age classes occurred which prevents
 286 the models from being used as estimates of coefficients for those age classes can not be obtained. In
 287 practice this could be avoided by reducing the number of age groups below where the low sampling
 288 numbers occur.

289 For each simulation the total stratified abundance at age is calculated. This follows the methods
 290 outlined in Smith, Somerton, 1981 that developed from stratified random sampling techniques
 291 described in greater detail in books like Cochran, 1977 and Lohr, 2009. The survey area is divided
 292 into N trawlable units and H strata, where N_h is the number of trawlable units in strata h . The
 293 true mean catch at age a (Y_{ah}) in survey in stratum h is found by

$$\bar{Y}_{ah} = \frac{\sum_{i=1}^{N_h} y_{ahi}}{N_h}, \quad (17)$$

294 where y_{ahi} is the survey catch at age in the i th unit. The total population estimate at age a is then

$$\hat{Y}_a = N \sum_{h=1}^H \frac{N_h}{N} \bar{Y}_{ah} \quad (18)$$

295 (Cochran, 1977; Lohr, 2009; Smith, Somerton, 1981).

296

297 2.6 Application

298 Atlantic cod and American Plaice are distributed throughout NAFO subdivision 3Ps, but during
299 most years abundances is highest at particular locations such as the Halibut Channel (Cod) or the
300 southeast slope of St. Pierre Bank (Plaice). The ALK methods discussed above were applied to
301 both Cod and American Plaice data from DFO’s multi-species bottom trawl RV survey of NAFO
302 subdivision 3Ps (Ings et al., 2019; M. J. Morgan et al., 2020). For the model based methods, the
303 same model formulations given in Equations 10, 15 & 16 were applied to both datasets along with
304 the empirical ALK. Cod data are from the start of DFO’s inshore/offshore survey in 1996 to 2018.
305 American Plaice data are limited until 2013 due to the lack of aged otoliths(M. J. Morgan et al.,
306 2020). Samples for otolith collection for both species were subject to length-stratified sampling and
307 the number of otoliths collected varies from year to year(Ings et al., 2019). Cod otolith collection
308 uses a sampling scheme requiring otoliths to be collected from five different areas around NAFO
309 subdivision 3PsIngs et al., 2019. However American Plaice collection simply requires otoliths to
310 be collected from the entire area. In 2006 the survey was unsuccessfully completed(Ings et al.,
311 2019). The survey areas for the two species are similar but not the same since the survey area for
312 cod does not include all of strata (Ings et al., 2019; M. J. Morgan et al., 2020). Since the study
313 areas differ in size, two different meshes were required for each of the applications. For cod a more
314 detailed boundary and higher density mesh for a more exact approximation was used, while the
315 Plaice analysis was performed using a less dense mesh with a less detailed boundary. All methods
316 were applied to each year of data independently of one another.

317 The mesh design can have an impact on the performance of the model. If a mesh is not
318 dense enough, the approximation may not work well. Mesh designs can sometimes also impact
319 the convergence of the model. Care should also be taken to ensure that all points that should be
320 outside the boundary, are in fact. Further details on how to create a mesh for a barrier model can
321 be seen in the referenced tutorial (**goodMesh**). As with the simulation in the previous section, age
322 frequencies were generated using the same four methods and then stratified estimates of abundance
323 were obtained.

324 3 Results

325 3.1 Simulation Study

326 The two spatial models performed similarly with the GFB model having a lower RMSE (between
327 the true abundance at age and the estimated abundance at age across the same survey) than the
328 traditional ALK in 76.9% of simulations and the GAM model 67.8% of simulations. Overall 40.5%
329 of the time the GFB model had the lowest error across all models and the GAM model 36.8% of
330 the time. The upper part of Figure 1 is a pairs plot of RMSE comparing each method against one
331 another, while the diagonal elements are density plots of the RMSE of each model. The bottom of
332 Figure 1 is a boxplot of the RMSE for each method. Overall the two spatial models yield tighter
333 bounds than the non-spatial models.

334 The bottom row of Figure 2 is the true simulated spatial distribution for fish aged 4,5 & 6 for
335 one simulation. The ages are somewhat overlapping in their spatial distribution but are centred in
336 different areas. The spatial ALKs are better able to discriminate between ages by considering the
337 location of the samples. This is evident in the top row of Figure 2 which shows the probability of
338 being a fish being age 4, 5 or 6 given a length of 40 cm across the simulated study area as obtained
339 by the GFB model. The probability of being in that age class increases when predicting over the
340 main bulk of that age class. The overall trend of the GAM approach is very similar to that of the

341 GFB model while also displaying evidence of undesired smoothing underneath landmasses. Both
342 the GFB and GAM yield a higher probability for 5 year olds south of the peninsula than might be
343 expected by the true proportions. This is due to the fact that the survey caught a larger share of
344 age 5s than for other age classes.

345 Each fish's length in the population is simulated when the population is generated and then
346 distributed spatially into the simulation grid cells such that the age and length for every fish in
347 the simulation is known. The true proportion at age in each of the 4833 simulation grid cells was
348 compared to the predicted proportion at age for each of the four methods. Figure 3a captures the
349 difference between the true proportion of an age and the predicted proportion from the GFB model
350 in each of the simulation grid cells for the same simulation used in Figure 2. If the model was
351 perfect then the entire map would be a single solid color representing zero difference.

352 Figure 3b is the same style of plot, except the predicted proportions come from the traditional
353 ALK. Results are very similar to those for the non-spatial CRL model. Compared to the GFB
354 method it has larger differences in proportion for the older ages. Younger ages have more of the
355 difference in age proportion spread out across the space than concentrated in a single region than
356 the GFB model.

357 [Figure 1 about here.]

358 [Figure 2 about here.]

359 [Figure 3 about here.]

360 [Figure 4 about here.]

361 [Table 2 about here.]

362 To assess how well the GFB model improves performance near a landmass the RMSE for the
363 160 simulation grid cells surrounding the bay was calculated. Figure 4 shows the median RMSE
364 across all simulations in each of those grid cells by age from each of the four methods. The median
365 RMSE is typically much lower in the two spatial methods than the two non-spatial ones. Table
366 2 shows the percentage of the 160 cells where either the GFB model or GAM model has a lower
367 median RMSE. For ages 3 and below the GAM model has a lower median RMSE for most cells but
368 for ages 4 and up the GFB model outperforms it. On average ages 4 and up make up over 74% of
369 the biomass available to the survey in those 160 grid cell suggesting the GFB model represents an
370 improvement on the majority of fish.

371 Overall the simulation showed that both spatial ALKs methods are capable of improving esti-
372 mates of abundance at age from a stratified random survey. This suggests that the age frequencies
373 created by the spatial ALKs are more accurate. For ages that make up a larger share of the abun-
374 dance like ages 4 through 7 the reduction in error is very noticeable as can be seen in the example
375 between the GFB model and traditional ALK in Figure 3 and around the landmass for all methods
376 in Figure 4. However for other age classes like one and two the differences can be minor.

377 **3.2 Application**

378 Atlantic cod and American Plaice are distributed throughout NAFO subdivision 3Ps, but during
379 most years abundances is highest at particular locations such as the Halibut Channel (Cod) or the
380 southeast slope of St. Pierre Bank (Plaice). The ALK methods discussed above were applied to
381 both Cod and American Plaice data from DFO's multi-species bottom trawl RV survey of NAFO

382 subdivision 3Ps (Ings et al., 2019; M. J. Morgan et al., 2020). For the model based methods, the
383 same model formulations given in Equations 10, 15 & 16 were applied to both datasets along with
384 the empirical ALK. Cod data are from the start of DFO’s inshore/offshore survey in 1996 to 2018.
385 American Plaice data are limited until 2013 due to the lack of aged otoliths(M. J. Morgan et al.,
386 2020). Samples for otolith collection for both species were subject to length-stratified sampling and
387 the number of otoliths collected varies from year to year(Ings et al., 2019). Cod otolith collection
388 uses a sampling scheme requiring otoliths to be collected from five different areas around NAFO
389 subdivision 3PsIngs et al., 2019. However American Plaice collection simply requires otoliths to
390 be collected from the entire area. In 2006 the survey was unsuccessfully completed(Ings et al.,
391 2019). The survey areas for the two species are similar but not the same since the survey area for
392 cod does not include all of strata (Ings et al., 2019; M. J. Morgan et al., 2020). Since the study
393 areas differ in size, two different meshes were required for each of the applications. For cod a more
394 detailed boundary and higher density mesh for a more exact approximation was used, while the
395 Plaice analysis was performed using a less dense mesh with a less detailed boundary. All methods
396 were applied to each year of data independently of one another.

397 The mesh design can have an impact on the performance of the model. If a mesh is not
398 dense enough, the approximation may not work well. Mesh designs can sometimes also impact
399 the convergence of the model. Care should also be taken to ensure that all points that should be
400 outside the boundary, are in fact. Further details on how to create a mesh for a barrier model can
401 be seen in the referenced tutorial (**goodMesh**). As with the simulation in the previous section, age
402 frequencies were generated using the same four methods and then stratified estimates of abundance
403 were obtained.

404 3.2.1 American Plaice

405 American Plaice are associated with fine substrates and both juveniles and adults frequently occur
406 in the same habitats (M. Morgan, 2000; Johnson, 2004). They do not conduct extensive annual
407 migrations (Johnson, 2004). The model based methods were fit from ages one to thirteen except
408 for a handful of years where no age one Plaice otoliths sampled. In those cases the models were
409 run on ages two through thirteen. Estimates of total abundance at age for American Plaice for
410 the four different method are shown in Figure 5. With the exception of age one plaice, the four
411 methods result in very similar estimates for the majority of the time series of total abundance at
412 age (obtained by aggregating across space using the Stratified Mean Method (SMM)). Closer to the
413 end of the time series there is a divergence for some age groups like 7 and 8 due to data sparsity.

414 Despite the similarity in aggregated abundance metrics across methods, when looking at the
415 spatial GAM and GFB results there are differences in the predicted probabilities. For example in
416 Figure 6 , the 2003 survey year the unconditional probabilities of an American Plaice being age 8
417 with a length of 35 cm were predicted across the space for both the GFB and GAM models. There
418 does appear to be evidence of the probabilities being smoothed underneath the peninsula and into
419 the bay in the GAM version that does not occur with the GFB models due to its support for
420 physical barriers. Both models make it clear that the probabilities are spatially varying and there
421 is a dependence on location. When looking at other ages not shown here, the two spatial methods
422 do not always agree, the GAM method will sometimes predict almost flat gradients across space
423 while the GFB for the same year will vary more across space. Figure 6 also showcases examples
424 of spatial ALK constructed from the GFB and GAM models at two different points. One at the
425 tip of Placentia Bay, and one in Fortune Bay. Each curve represents the proportion at each length
426 taken up by that age (like the columns of an ALK) while each vertical slice at a length must sum
427 to one (like the rows of an ALK). Since age 13 was used as a plus group as fish get longer they end

428 up more likely to land there. The two ALKs are different at each of the two points. For instance,
429 the ALKs estimated by the GFB model shows more overlap between ages 1 and 2 in Placentia
430 Bay than Fortune Bay while still having more overlap between the first two ages when compared
431 against the ALKs at the same points from the GAM model.

432 [Figure 5 about here.]

433 [Figure 6 about here.]

434 3.2.2 Cod

435 In contrast to American Plaice, Atlantic Cod occur over a broader range of substrates with juveniles
436 and adults overlapping in some broad areas whereas only juveniles may be more frequently sampled
437 at some locations closer to the coast(Dalley, Anderson, 1997; Fahay et al., 1999). Adult Atlantic
438 Cod conduct extensive migrations from areas offshore to shallow coastal locations and there is
439 evidence to suggest some alongshore movements as well(Fahay et al., 1999; Brattey et al., 2002).

440 For the Cod data, age eleven was taken to be the plus group for every year and the models
441 were fit to ages one through eleven for all years. As for the American Plaice data the stratified
442 estimates of abundance were generated using all four methods and are shown in Figure 7. Each
443 of the four aging methods result in similar trends. The two non-spatial methods both show very
444 similar trends with almost completely overlapping lines in most years. The spatial methods, GAM
445 and GFB are largely similar in trend but differ occasionally from the non-spatial methods in (e.g.
446 age 10 and 11 in 2004).

447 ALKs were generated at two different locations within 3Ps. Figure 8 displays the locations
448 used to create these ALKs using 3 cm length bins along with corresponding visual representations.
449 ALKs constructed at each of the two areas are different from one another. The GAM model finds
450 proportion of fish being age 10 and 11 in Fortune Bay to be essentially zero and a similar situation
451 occurs in Placentia Bay except for the last few age groups.

452 The unconditional probabilities of Cod being a certain age given various lengths were examined
453 spatially for both the GFB and GAM methods. Both of the spatial methods suggest that there is
454 a difference spatially in age for Cod of a given length. An example of this for the probabilities of
455 being age 5 for 50 cm Cod can be seen in the left hand side of Figure 8 for both the GAM and
456 GFB models. The GFB finds a lower probability of Cod being age 5 in the tip of Fortune Bay,
457 the GAM model however shows evidence of smoothing underneath the landmass.Using the GFB
458 model it can be seen that Cod on the northwestern portion of the survey area also have a much
459 lower probability of being age 5 at 50 cm than those that live in the rest of survey area.

460 The simulation study presented in Section 2.5 was designed to have an age and length structure
461 that mimics the 3Ps cod population with a stratified survey design similar to the one used in the
462 region. Based on the results of the simulation study it is expected that for most years the GFB
463 model should provide more accurate estimates of the abundance at age for ages 4 and up.

464 [Figure 7 about here.]

465 [Figure 8 about here.]

466 4 Discussion

467 Errors related to the process of obtaining age estimates are often ignored in age structured stock
468 assessment models. Model based approaches offer an avenue for incorporating such errors even

469 when a non-spatial method of aging is used. If errors in the aging process are ignored estimates
470 obtained from the stock assessment model provide a false sense of precision and may impact derived
471 quantities (e.g. spawning stock biomass). The simulation study here also suggests that spatial
472 methods have the potential to further reduce errors resulting from applying an ALK. Future work
473 includes incorporating a spatial ALK model into a stock assessment model directly to see exactly
474 how derived quantities and associated errors may be affected by the age estimation process.

475 Incorporating more spatial information into the stock assessment process has the potential
476 to increase precision. It is still not uncommon for stock assessment models to simply ignore or
477 aggregate over space (Punt, 2019), often taking an areal approach instead of a pointwise one. Our
478 GFB model offers another choice of spatial ALK that could be integrated into stock assessment
479 models particularly when there may be landmasses present in the study area.

480 In addition, both the simulation study and the application suggest that even in cases where it
481 is not possible or desirable to fit a spatial ALK model due to data limitations or other constraints
482 it may still be worthwhile using a model based approach to construct ALKs in order to gain the
483 aforementioned benefits of smoothing and bridging of gaps.

484 Spatial ALKs can provide improved estimates at age over non-spatial methods. The simulation
485 study showed that over three quarters of the time using a spatial method to generate the ALK had
486 a reduced RMSE for the true abundance numbers at age as compared to traditional methods. It
487 also suggested that spatial methods can reduce the error across all ages in both the entire study
488 area, and in spatial pockets for most ages. Examining the probabilities of a fish being a given
489 length may indeed give insight into where certain age classes may be distributed during the time
490 of the survey.

491 The GFB had the lowest RMSE more often than the other three models indicating it led to
492 abundance at age estimates closer to the true values. Since the simulations distributed the fish
493 randomly among the survey grid, in some simulations the majority of fish may not have been close
494 to the landmass limiting the performance benefits of the GFB model to be similar to the GAM
495 model. The GFB model also had the lowest maximum RMSE among all models suggesting it will
496 not perform worse in most scenarios. The GFB model also provides more realistic plots of the
497 probabilities of age given length by preventing them from smoothing beneath landmasses.

498 When looking at the two applications presented, neither of the spatial methods made obviously
499 large changes to the abundance indices at age. However when examining the predicted probabilities
500 at age there is a clear indication that they do vary spatially. While there are similarities between
501 the two spatial methods for the bulk of the probabilities predicted, the fact that the GFB model
502 supports physical barriers is evident in how the bay is treated as captured in Figures 6 & 8. ALKs
503 are demonstrated to vary with space.

504 This work makes evident that applying a non-spatial ALK may have ramifications for calculating
505 indices of abundance. Combined with the fact that errors from the aging process are often ignored,
506 there is a strong argument for integrating spatial ALK models directly into stock assessment models.

507 **References**

- 508 [1] S. Aanes, J. H. Vølstad, Efficient statistical estimators and sampling strategies for estimating
509 the age composition of fish, *Canadian Journal of Fisheries and Aquatic Sciences* 72 (6) (2015)
510 938–953, DOI: 10.1139/cjfas-2014-0408.
- 511 [2] W. H. Aeberhard, J. Mills Flemming, A. Nielsen, Review of state-space models for fisheries
512 science, *Annual Review of Statistics and Its Application* 5 (2018) 215–235, DOI: 10.1146/
513 *annurev-statistics-031017-100427*.

- 514 [3] A. Agresti, *Categorical data analysis*, volume 482, John Wiley & Sons, 2003, DOI: 10.1002/
515 0471249688.
- 516 [4] H. Bakka, How to solve the stochastic partial differential equation that gives a Matérn random
517 field using the finite element method, Mar. 2018.
- 518 [5] H. Bakka, H. Rue, et al., Spatial modelling with R-INLA: A review, *Wiley Interdisciplinary
519 Reviews: Computational Statistics* 10 (6) (2018).
- 520 [6] H. Bakka, J. Vanhatalo, J. Illian, et al., Accounting for physical barriers in species distribution
521 modeling with non-stationary spatial random effects, arXiv preprint arXiv:1608.03787 (2016).
- 522 [7] H. Bakka, J. Vanhatalo, J. B. Illian, et al., Non-stationary Gaussian models with physical
523 barriers, *Spatial Statistics* (2019), DOI: 10.1016/j.spasta.2019.01.002.
- 524 [8] C. W. Berg, K. Kristensen, Spatial age-length key modelling using continuation ratio logits,
525 *Fisheries Research* 129 (2012) 119–126, DOI: 10.1016/j.fishres.2012.06.016.
- 526 [9] J. Bratley, D. Porter, C. George, Exploitation rates and movements of Atlantic cod (*Gadus
527 morhua*) in NAFO Subdiv. 3Ps based on tagging experiments conducted during 1997 (2002).
- 528 [10] W. G. Cochran, *Sampling Techniques*, John Wiley & Sons Inc., 1977.
- 529 [11] E. Dalley, J. Anderson, Age-dependent distribution of demersal juvenile Atlantic cod (*Gadus
530 morhua*) in inshore/offshore northeast Newfoundland, *Canadian Journal of Fisheries and
531 Aquatic Sciences* 54 (S1) (1997) 168–176, DOI: 10.1139/f96-171.
- 532 [12] M. P. Fahay et al., Essential fish habitat source document. Atlantic cod, *Gadus morhua*, life
533 history and habitat characteristics (1999).
- 534 [13] A. Fridriksson, On the calculation of age-distribution within a stock of cod by means of
535 relatively few age-determinations as a key to measurements on a large scale, *Rapports Et
536 Proces-Verbaux Des Reunions, Conseil International Pour l’Exploration De La Mer* 86 (1934)
537 1–5.
- 538 [14] R. J. Gray, Flexible methods for analyzing survival data using splines, with applications to
539 breast cancer prognosis, *Journal of the American Statistical Association* 87 (420) (1992) 942–
540 951, DOI: 10.1080/01621459.1992.10476248.
- 541 [15] F. E. Harrell, *Regression modeling strategies, as implemented in R package ‘rms’ version 3* (3)
542 (2014), DOI: 10.1007/978-1-4757-3462-1.
- 543 [16] R. Hilborn, D. Ovando, Reflections on the success of traditional fisheries management, *ICES
544 journal of Marine Science* 71 (5) (2014) 1040–1046, DOI: 10.1093/icesjms/fsu034.
- 545 [17] D. Ings et al., Assessing the status of the cod (*Gadus morhua*) stock in NAFO Subdivision
546 3Ps in 2018, *Fisheries & Oceans Canada, Science, Canadian Science Advisory Secretariat*,
547 2019.
- 548 [18] D. Johnson, American plaice, *Hippoglossoides platessoides*, life history and habitat charac-
549 teristics, *NOAA Tech. Mem. NMFS-NE* 187 (2004).
- 550 [19] K. Kristensen et al., TMB: Automatic differentiation and laplace approximation, *Journal of
551 Statistical Software* 70 (5) (2016) 1–21, DOI: 10.18637/jss.v070.i05.
- 552 [20] T. Kvist, H. Gislason, P. Thyregod, Using continuation-ratio logits to analyze the variation
553 of the age composition of fish catches, *Journal of applied statistics* 27 (3) (2000) 303–319,
554 DOI: 10.1080/02664760021628.

- 555 [21] F. Lindgren, H. Rue, An explicit link between Gaussian fields and Gaussian Markov random
556 fields: the stochastic partial differential equation approach, *Journal of the Royal Statistical*
557 *Society. Series B* (2011), DOI: 10.1111/j.1467-9868.2011.00777.x.
- 558 [22] S. L. Lohr, *Sampling: design and analysis*, Nelson Education, 2009, DOI: 10.1201/9780429296284.
- 559 [23] M. J. Morgan et al., Assessing the status of the American Plaice (*Hippoglossoides plates-*
560 *soides*) stock in NAFO Subdivision 3Ps in 2019, Fisheries & Oceans Canada, Science, Cana-
561 dian Science Advisory Secretariat, 2020.
- 562 [24] M. Morgan, Interactions between substrate and temperature preference in adult American
563 plaice (*Hippoglossoides platessoides*), *Marine & Freshwater Behaviour & Phy* 33 (4) (2000)
564 249–259, DOI: 10.1080/10236240009387096.
- 565 [25] J. K. Parrish, Using behavior and ecology to exploit schooling fishes, *Environmental Biology*
566 *of Fishes* 55 (1-2) (1999) 157–181.
- 567 [26] A. E. Punt, Modelling recruitment in a spatial context: A review of current approaches,
568 simulation evaluation of options, and suggestions for best practices, *Fisheries Research* 217
569 (2019) 140–155, DOI: 10.1016/j.fishres.2017.08.021.
- 570 [27] A. E. Punt, M. Haddon, G. N. Tuck, Which assessment configurations perform best in the
571 face of spatial heterogeneity in fishing mortality, growth and recruitment? A case study based
572 on pink ling in Australia, *Fisheries research* 168 (2015) 85–99, DOI: 10.1016/j.fishres.
573 2015.04.002.
- 574 [28] P. M. Regular et al., SimSurvey: An R package for comparing the design and analysis of
575 surveys by simulating spatially-correlated populations, *PloS one* 15 (5) (2020) e0232822, DOI:
576 10.1371/journal.pone.0232822.
- 577 [29] A. Rindorf, P. Lewy, Analyses of length and age distributions using continuation-ratio logits,
578 *Canadian Journal of Fisheries and Aquatic Sciences* 58 (6) (2001) 1141–1152, DOI: 10.1139/
579 f01-062.
- 580 [30] S. M. Ross, *Introduction to probability models*, Academic press, 2014.
- 581 [31] H. Rue, L. Held, *Gaussian Markov random fields: theory and applications*, CRC press, 2005,
582 DOI: 10.1201/9780203492024.
- 583 [32] S. Smith, G. Somerton, STRAP: A User-Oriented Computer Analysis System for Groundfish
584 Research Trawl Survey Data, technical report, Department of Fisheries and Oceans, Sept.
585 1981.
- 586 [33] T. Stari et al., Smooth age length keys: observations and implications for data collection on
587 North Sea haddock, *Fisheries Research* 105 (1) (2010) 2–12, DOI: 10.1016/j.fishres.2010.
588 02.004.
- 589 [34] J. T. Thorson, L. A. Barnett, Comparing estimates of abundance trends and distribution
590 shifts using single-and multispecies models of fishes and biogenic habitat, *ICES Journal of*
591 *Marine Science* 74 (5) (2017) 1311–1321, DOI: 10.1093/icesjms/fsw193.
- 592 [35] J. T. Thorson, A. O. Shelton, et al., Geostatistical delta-generalized linear mixed models
593 improve precision for estimated abundance indices for West Coast groundfishes, *ICES Journal*
594 *of Marine Science* 72 (5) (2015) 1297–1310, DOI: 10.1093/icesjms/fsu243.
- 595 [36] P. J. Verweij, H. C. Van Houwelingen, Penalized likelihood in Cox regression, *Statistics in*
596 *medicine* 13 (23-24) (1994) 2427–2436, DOI: 10.1002/sim.4780132307.

597 [37] S. N. Wood, M. V. Bravington, S. L. Hedley, Soap film smoothing, *Journal of the Royal*
598 *Statistical Society: Series B (Statistical Methodology)* 70 (5) (2008) 931–955, DOI: 10.1111/
599 *j.1467-9868.2008.00665.x*.

600 [38] B. Worm et al., Rebuilding global fisheries, *science* 325 (5940) (2009) 578–585, DOI: 10.1126/
601 *science.1173146*.

602 Appendices

603 A Simulation Code

604 The simulation study uses largely the same process for `SimSurvey` as was outlined and described
605 in Regular et al., 2020. To assess the effectiveness of each aging method spatially (as described in
606 Figure 3) the lengths and ages of every single fish needs to be simulated. `SimSurvey` only applies
607 individual lengths as fish are sampled, only storing the number of fish of each age in the simulation
608 grid cells which greatly reduces storage costs. To have the lengths available for every single fish
609 `SimSurvey` was modified to generate fish lengths from the growth curve at the same time as the
610 population is generated. These lengths are then kept as fish are distributed spatially into the
611 simulation grid cells. The modified functions allowing all individual fish lengths to be known were
612 created in a private fork of the `SimSurvey` package.

613 In addition the Age-year-space covariance discussed in Appendix S3 of Regular et al., 2020 was
614 instead obtained using a GMRF approximation with support for physical barriers as described in
615 Section 2.2.1. A precision matrix Q is generated from a mesh and a specified set of hyperparameters.
616 This is then approximated for each grid point in the `SimSurvey` simulation. This has the benefit
617 over the default `SimSurvey` method of constraining the simulation to also have to abide by any
618 physical barriers present and also provides a speed boost by using a GMRF approximation when the
619 mesh has less nodes than there are cells in the grid. An example of a mesh used in the simulation
620 is shown in Figure 9 and example of the true abundance at age for the simulation referred to in
621 Figures 3a and 2 is shown in Figure 10.

622 [Figure 9 about here.]

623 [Figure 10 about here.]

624 [Figure 11 about here.]

625 [Figure 12 about here.]

626 [Figure 13 about here.]

627 [Figure 14 about here.]

628 [Figure 15 about here.]

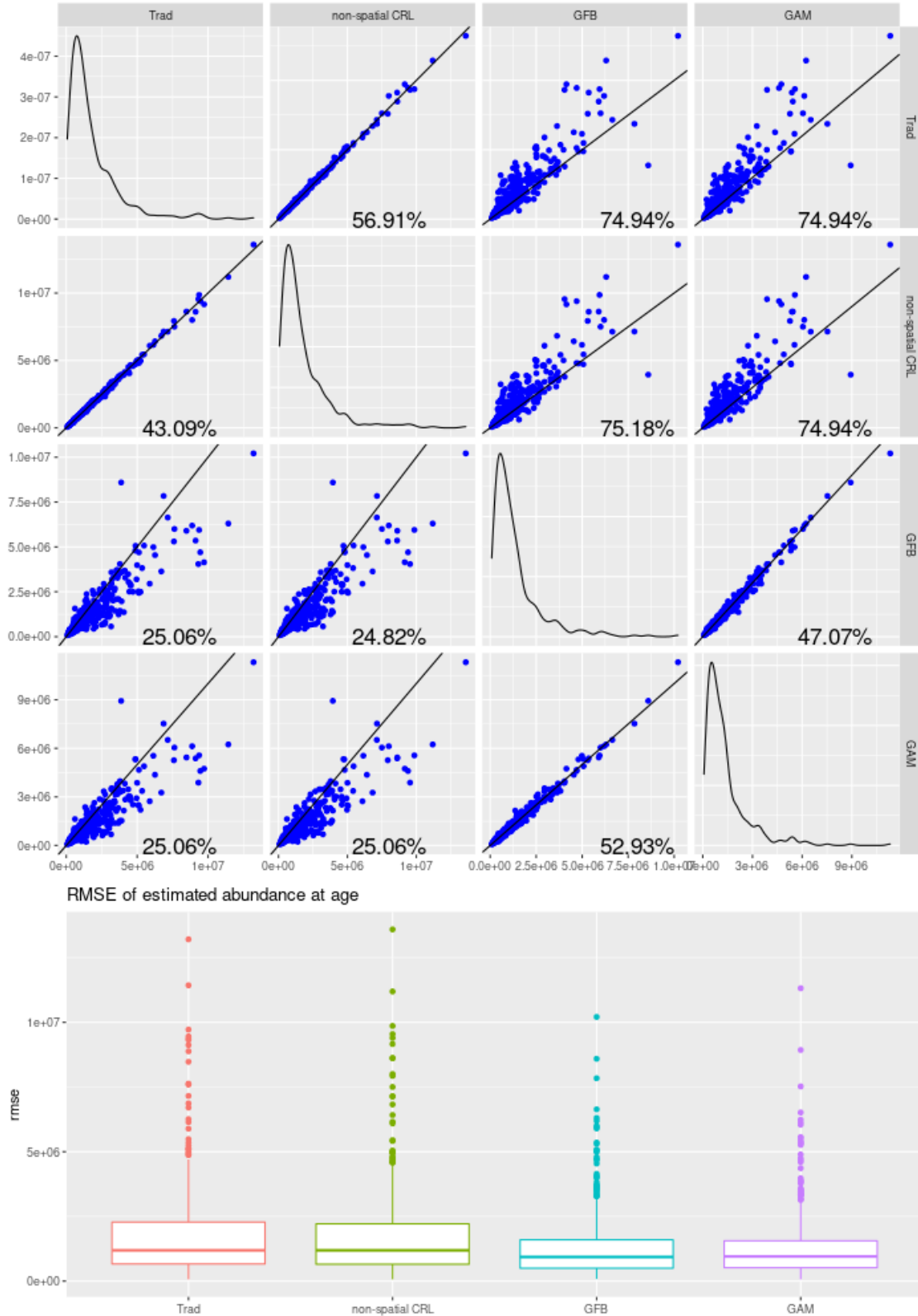


Figure 1: The RMSE of the stratified survey estimate of abundance at age versus the true total abundance at age in the simulation. The top figure compares the RMSE of each of the methods against one another with a diagonal line illustrates where points would lie if the two methods were identical. The percentages are the percent of simulations where the RMSE of the model on the x-axis is less than or equal to the RMSE of the model on the y-axis. Plots along the diagonal of the top figure are density plots of the RMSE for each method. The bottom part of the figure is a boxplot of the RMSE for each method.

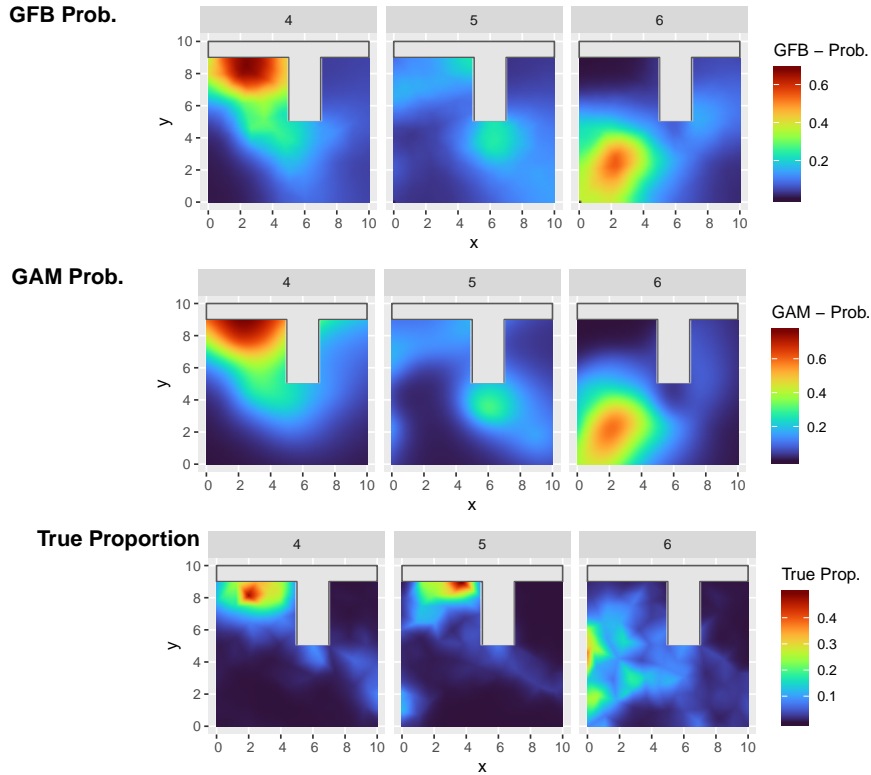
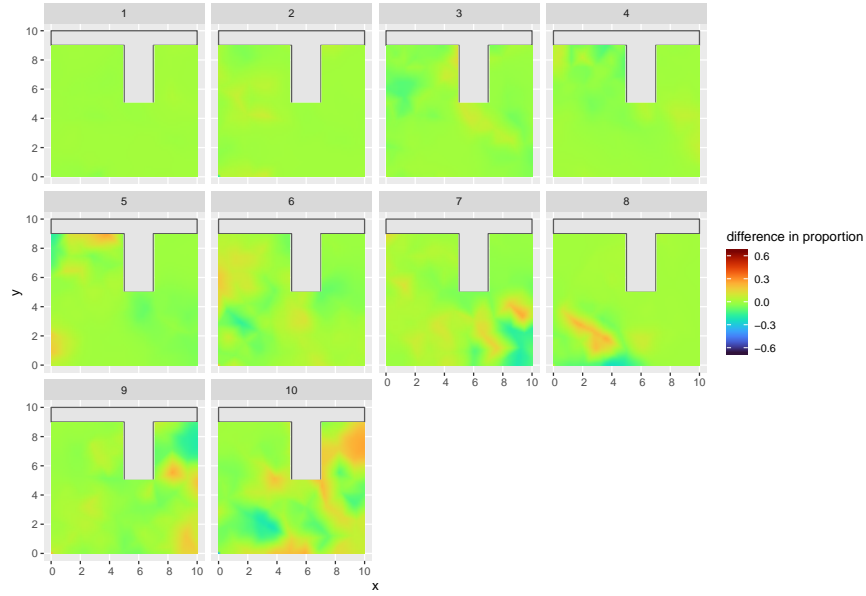
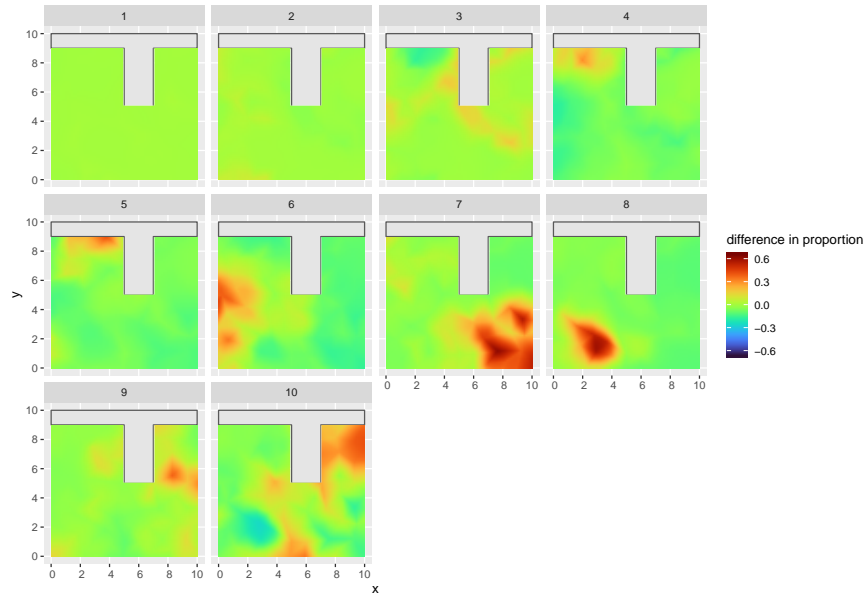


Figure 2: Top row is predicted probabilities in each grid cell of fish being aged 4,5, and 6 in one simulation with a length of 40 cm as predicted by the GFB model. Middle is the predicted probabilities for the same as predicted by the GAM model and the bottom row is the true simulated proportion of fish aged 4,5 and 6 in each grid cell as distributed by SimSurvey. The non-spatial CRL model found the probability of 23.1%, 10.9% and 18.6% for fish being aged 4, 5 and 6 respectively having a length of 40 cm. The traditional ALK method found the probability of 27.8%, 5.6% and 22.2% for fish being aged 4,5 and 6 respectively.

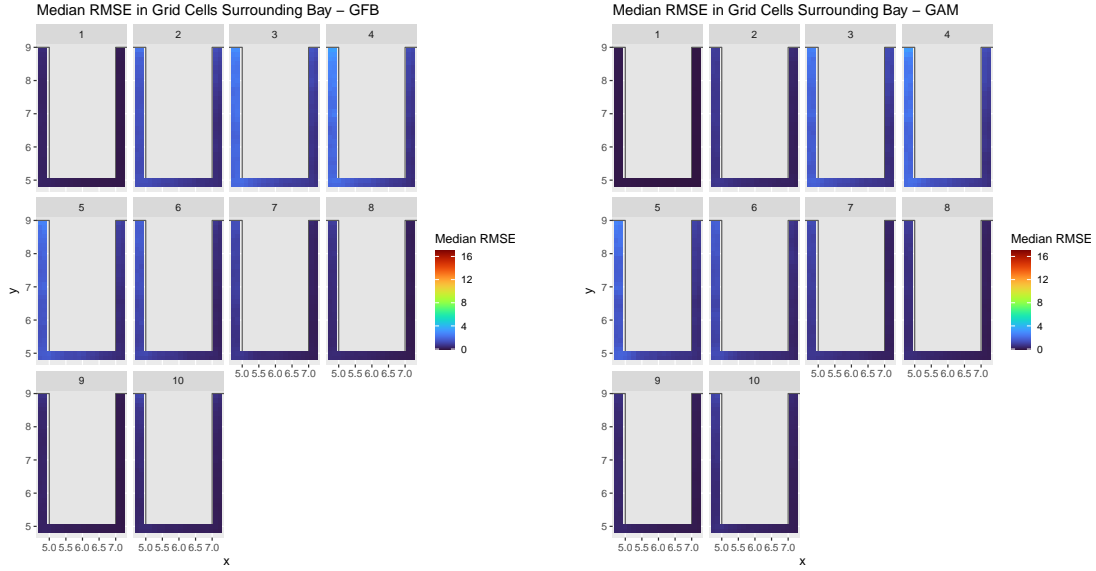


(a) Estimation error (true minus predicted) for the GFB model. A new spatial ALK is generated and applied to each simulation grid cell.



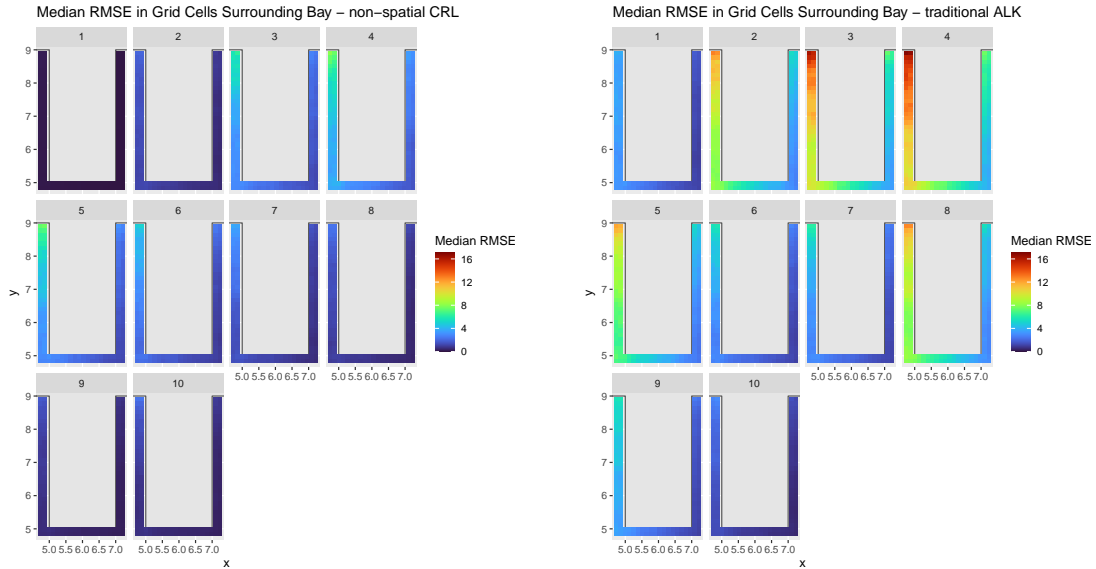
(b) Estimation error (true minus predicted) from the traditional ALK. The same global traditional ALK was applied to each simulation grid cell. The non-spatial CRL model results in a very similar plot to the traditional ALK. The traditional ALK struggles more with older ages than the GFB model.

Figure 3: The difference between the true proportion for each age in every grid cell and the proportion predicted by GFB model and traditional ALK for the same simulation as in Figure 2 for fish available to the survey. A perfect model would be a flat green with no difference in proportion between the two.



(a) GFB

(b) GAM



(c) non-spatial CRL

(d) traditional ALK

Figure 4: The median RMSE across all the simulations for the 160 grid cells surrounding the bay for each of the four methods by age. Both spatial models perform considerably better in each cell than the non-spatial methods. The x and y are the spatial coordinates used in the simulation and grey is the landmass.



Figure 5: Total abundance estimates by age for American Plaice within NAFO division 3P for the years from 1996 to 2013, not including 2006. Abundance estimates at age are largely similar across all four methods and follow the same general trends.

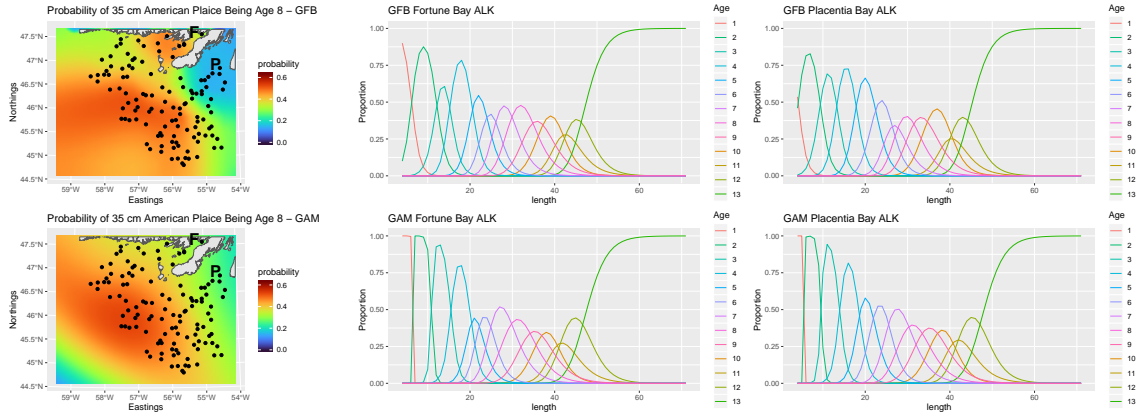


Figure 6: A visual representation of three ALKs generated using the GF & GAM methods at the two locations shown (P-Placentia Bay, F-Fortune Bay) for the 2003 survey year. Each vertical slice of the graphs in columns 2 and 3 must sum to 1 and is like a row in an ALK. The maps are also surfaces showing the predicted probabilities of American Plaice being 8 years old with a length of 35 cm. ALKs constructed at different locations result in different ALKs with either model. Points represent sampling locations from which otoliths were collected during that year.

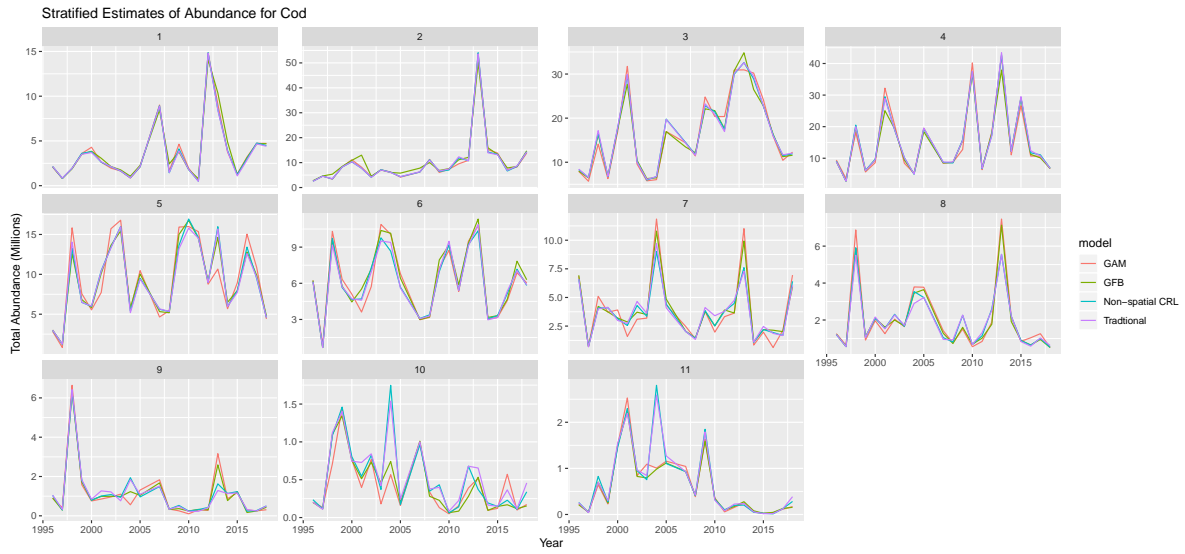


Figure 7: Total abundance estimates by age for the cod dataset for the years from 1996 to 2018, not including 2006. Trends across years are broadly similar between methods.

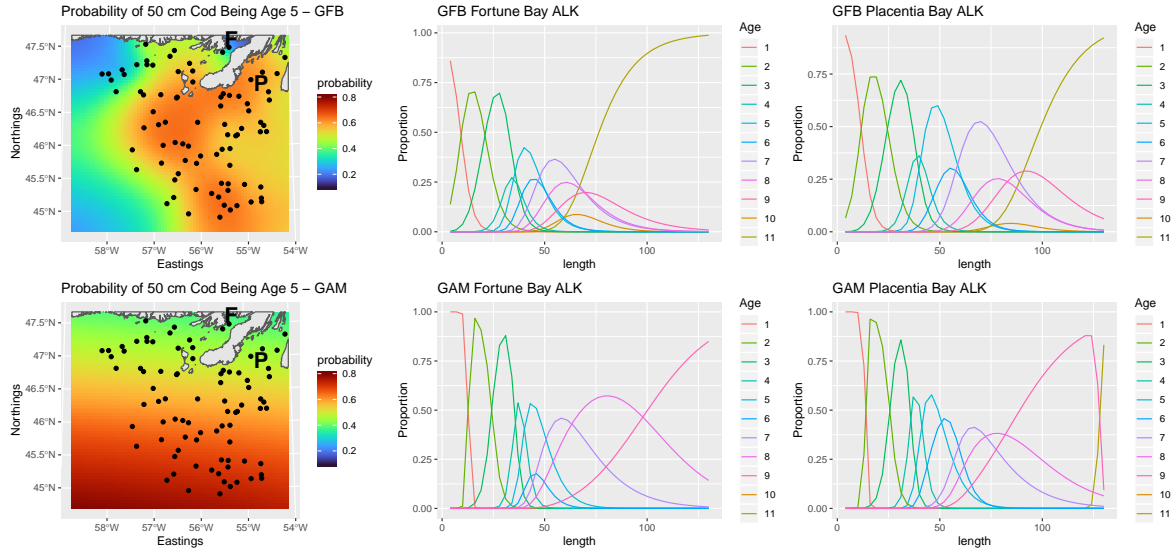


Figure 8: A visual representation of three ALKs generated using the GF & GAM methods at the two locations shown (P-Placentia Bay, F-Fortune Bay) for the 2011 survey year. Each vertical slice of the graphs in columns 2 and 3 must sum to 1 and is like a row in an ALK. The maps are also surfaces showing the predicted probabilities of cod being 5 years old with a length of 50 cm. ALKs constructed at different locations result in different ALKs with either model. Points represent sampling locations from which otoliths were collected during that year.

Constrained refined Delaunay triangulation

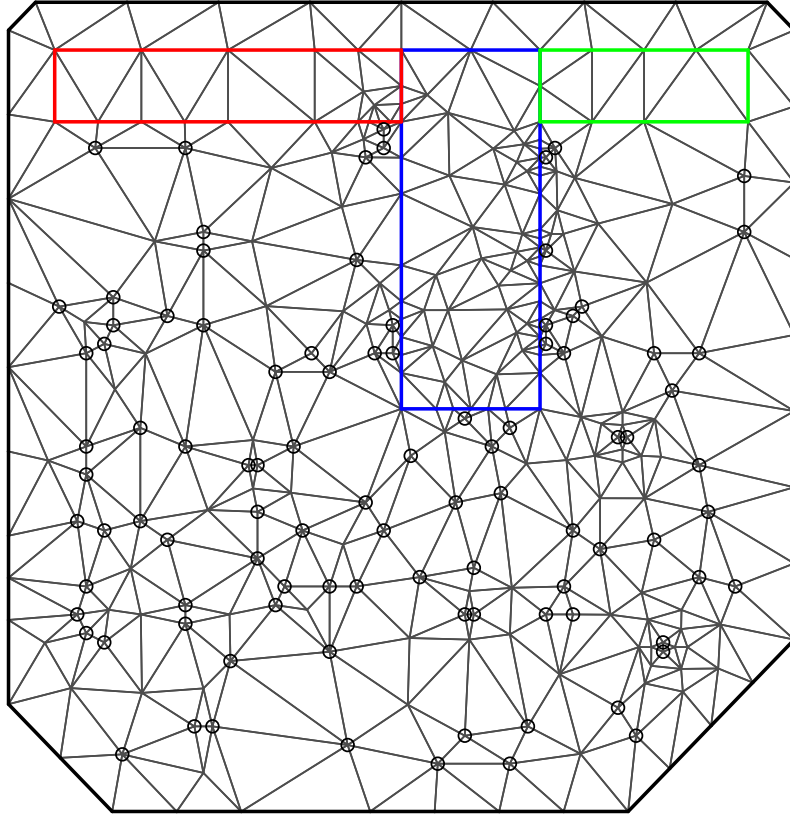


Figure 9: An example of a mesh of the simulated survey area for one simulation. The coloured rectangles make up the boundary area, circles are sampled locations. Triangles form the mesh and each vertex is an element in the covariance matrix.

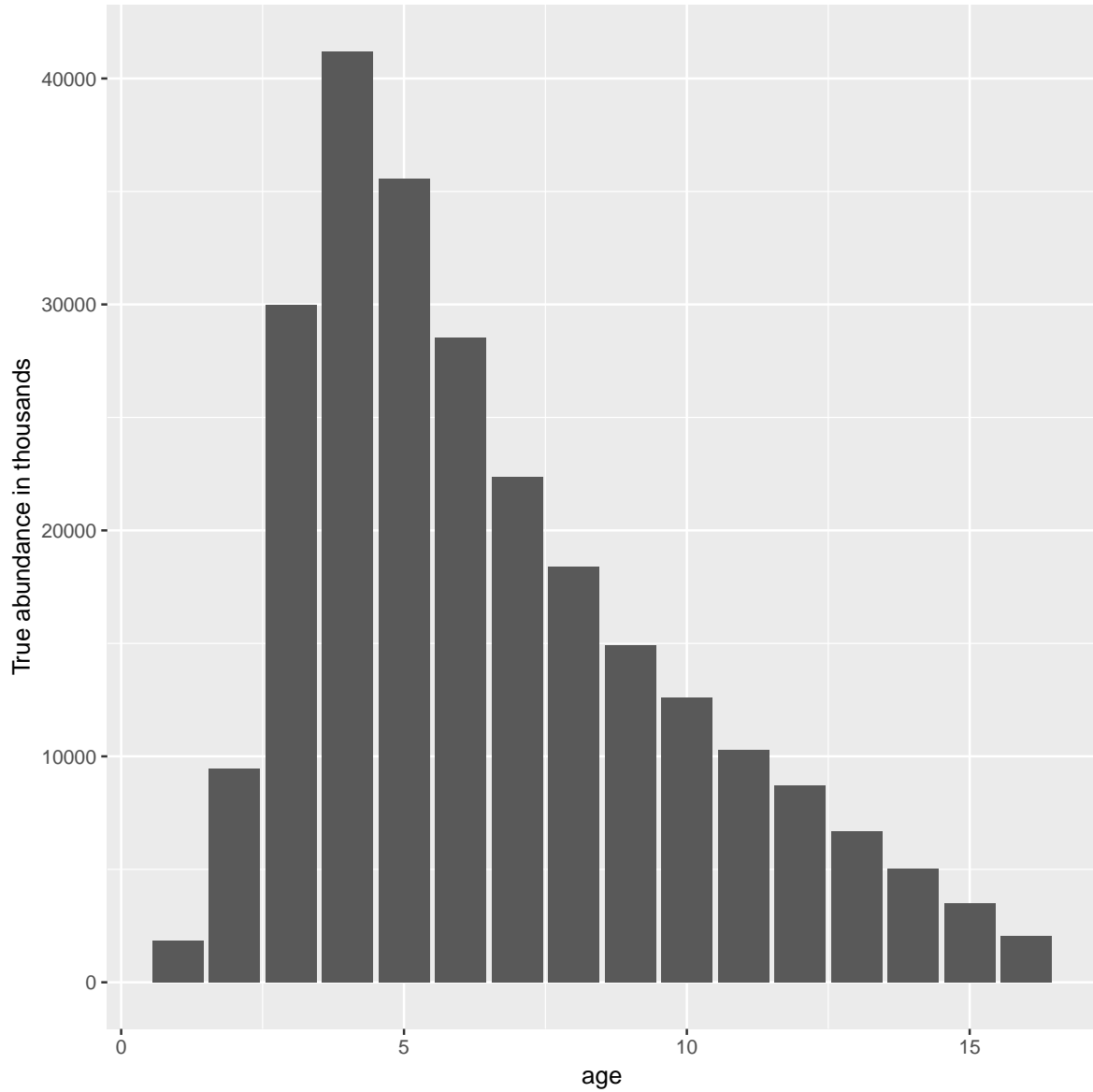


Figure 10: The true abundance at age (in thousands) for the simulation used in Figures 3 & 2

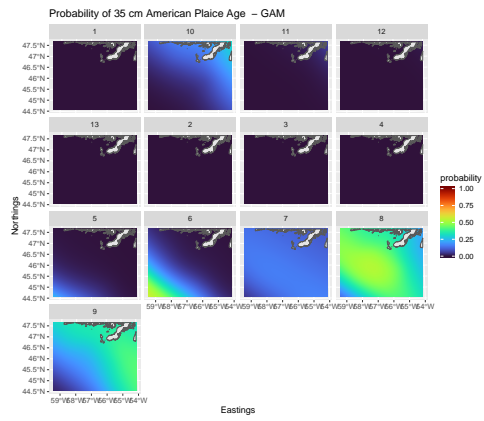


Figure 11: The Probability of an American Plaice being each age class in the study area given a length of 35 cm as predicted by the GAM model.

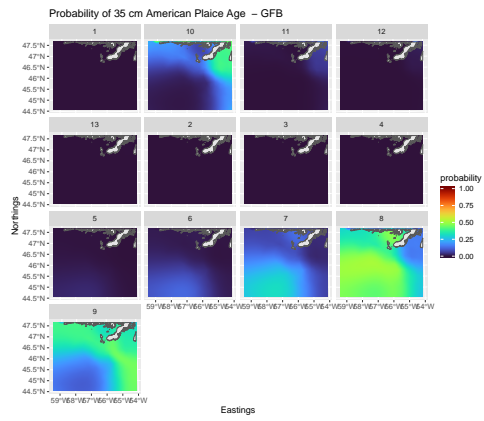


Figure 12: The Probability of an American Plaice being each age class in the study area given a length of 35 cm as predicted by the GFB model.

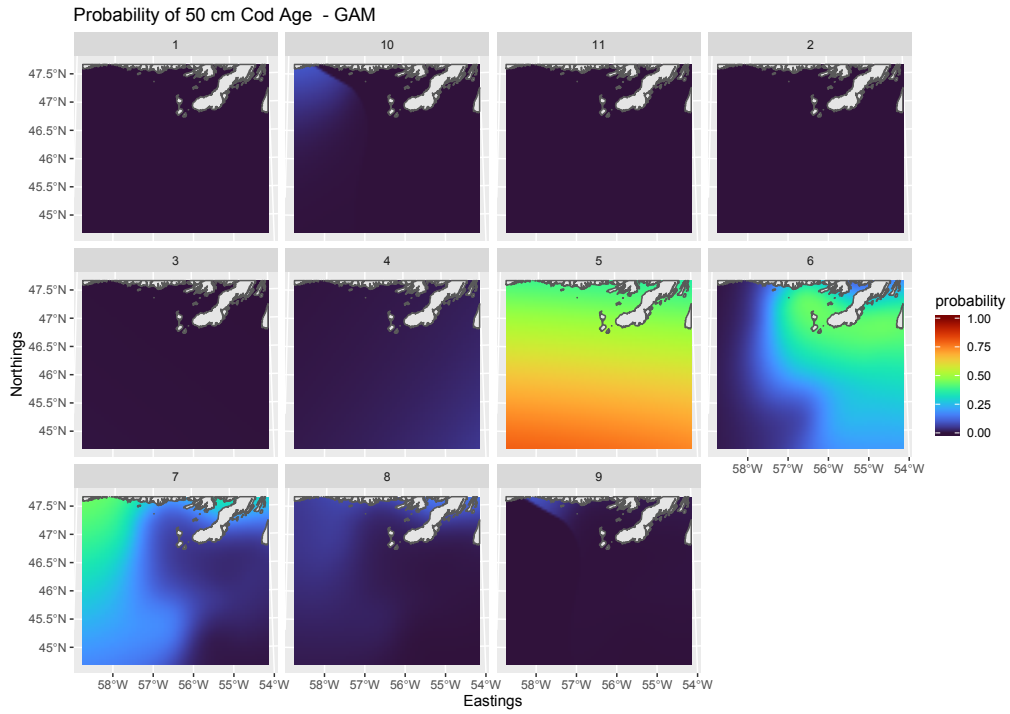


Figure 13: The Probability of a Cod being each age class in the study area given a length of 50 cm as predicted by the GAM model.

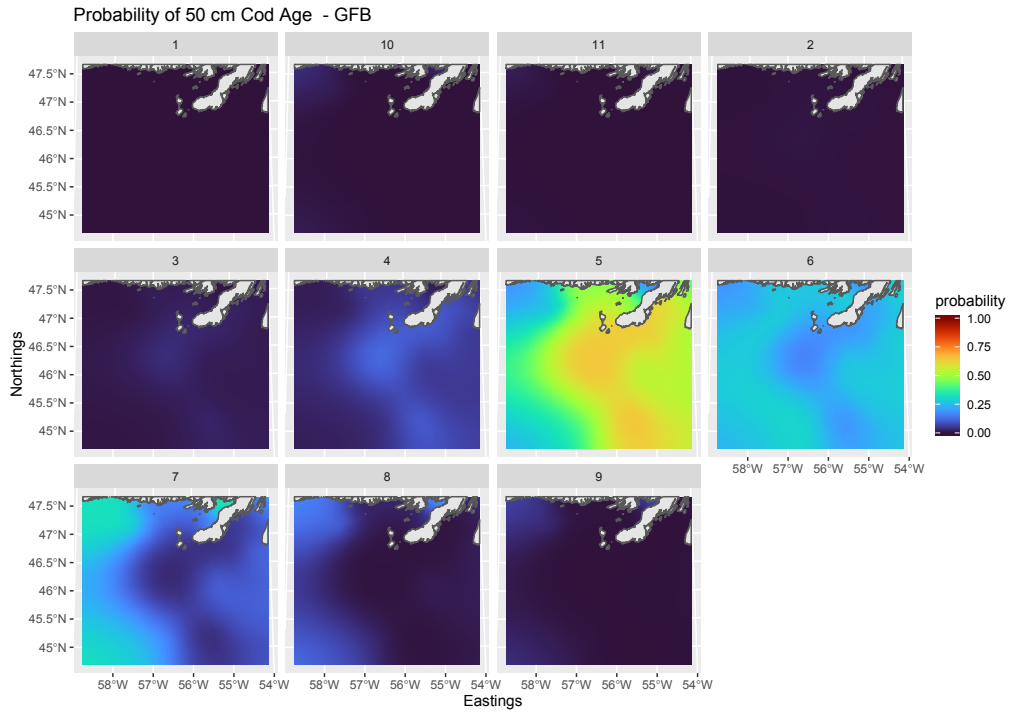
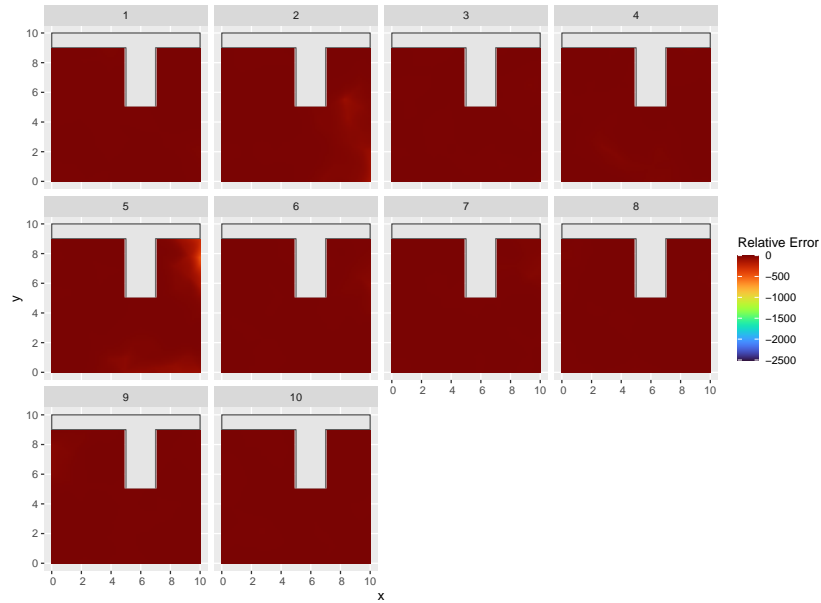
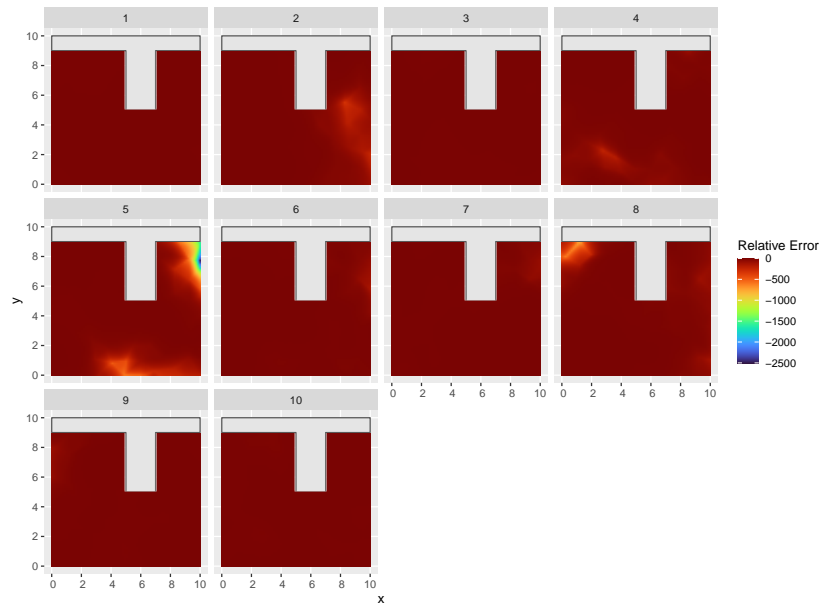


Figure 14: The Probability of a Cod being each age class in the study area given a length of 50 cm as predicted by the GFB model.



(a) Relative error (true minus predicted over true) for the GFB model. A new spatial ALK is generated and applied to each simulation grid cell.



(b) Relative error (true minus predicted over true) from the traditional ALK. The same global traditional ALK was applied to each simulation grid cell. The non-spatial CRL model results in a very similar plot to the traditional ALK.

Figure 15: The relative error between the predicted proportions and the true proportions in each simulation cell. The median relative error for the GFB model is -0.1334 and -0.4038 for the traditional ALK. Both methods have the smallest relative error occur on the right hand side for age 5 fish which is -467.9 for the GFB model and -2443.91 for the traditional ALK.

(a) A traditional ALK. Any fish measured to be less than 4 cm or greater than 52 cm would be missed by this ALK.

	1	2	3	4	5	6	7	8	9	10
1										
4										
7	1.000									
10	0.933	0.067								
13	0.200	0.767	0.033							
16		0.567	0.433							
19		0.100	0.700	0.200						
22			0.333	0.533	0.133					
25			0.033	0.433	0.467	0.067				
28				0.200	0.400	0.367	0.033			
31				0.033	0.167	0.467	0.333			
34					0.033	0.367	0.433	0.100	0.067	
37					0.067	0.100	0.400	0.267	0.167	
40						0.036	0.107	0.393	0.357	0.107
43						0.071		0.357	0.500	0.071
46								0.286	0.429	0.286
49									1.000	
52										
55										

(b) An example of a smooth ALK. This ALK was constructed from a CRL model with length as the sole covariate using the same age-growth data as was used to construct the ALK in Table 1a. All lengths are represented and there is no longer any possibility of missing fish at the more extreme length bins. Effects of sampling artifacts like all 49cm fish being considered age 9 despite shorter fish being age 10 are reduced. Zeros are again omitted for readability.

	1	2	3	4	5	6	7	8	9	10
1	1.000									
4	1.000									
7	1.000									
10	0.990	0.010								
13	0.064	0.918	0.018							
16		0.612	0.382	0.006						
19		0.047	0.798	0.150	0.004					
22		0.002	0.290	0.602	0.101	0.005				
25			0.031	0.473	0.408	0.083	0.005			
28			0.003	0.139	0.455	0.337	0.065	0.001		
31				0.027	0.209	0.476	0.273	0.010	0.005	
34				0.005	0.062	0.326	0.475	0.081	0.048	0.003
37				0.001	0.016	0.146	0.346	0.265	0.203	0.023
40					0.004	0.053	0.115	0.383	0.378	0.067
43					0.001	0.018	0.026	0.370	0.458	0.127
46						0.006	0.005	0.314	0.471	0.203
49						0.002	0.001	0.255	0.443	0.299
52						0.001		0.201	0.389	0.409
55								0.157	0.319	0.524

Table 1: Examples of traditional and smoothed ALKs. Columns represent ages and rows define three centimetre length bins. Zeros are omitted for readability.

Table 2: The percentage of the 160 grid cells surrounding the simulation bay where the median RMSE in each cell from all simulations is lower for either the GAM or GFB models by age. For ages four and up the GFB model outperforms the GAM which is the majority of biomass available to the survey.

	1	2	3	4	5	6	7	8	9	10
GAM	100.000	100.000	92.500	38.125	28.750	28.750	31.875	10.000	1.875	41.250
GFB	0.000	0.000	7.500	61.875	71.250	71.250	68.125	90.000	98.125	58.750