

# A gaussian field approach to generating spatial age length keys

Jonathan Babyn<sup>a,\*</sup>, Divya Varkey<sup>b</sup>, Paul Regular<sup>b</sup>, Danny Ings<sup>b</sup>, Joanna Mills Flemming<sup>a</sup>

<sup>a</sup> Dalhousie University, Department of Mathematics & Statistics, 6316 Coburg Road, PO Box 15000, Halifax NS B3H 4R2, Canada

<sup>b</sup> Fisheries and Oceans Canada, Northwest Atlantic Fisheries Center, St. John's, Newfoundland and Labrador, Canada

## ARTICLE INFO

Handled by Prof. A.E. Punt

### Keywords:

Spatial age-length keys  
Smooth age-length keys  
Age estimates

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

ALKs have been used to estimate the age of fish for over 80 years

(Fridriksson, 1934). They are based on the idea that the proportion of fish at age  $a$ ,  $p_a$  is equal to

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

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

Smooth ALKs have non-zero proportions at every possible length bin ensuring that all fish are assigned an age estimate. Kvist et al. (2000)

\* Corresponding author.

E-mail addresses: [jn805248@dal.ca](mailto:jn805248@dal.ca) (J. Babyn), [Divya.Varkey@dfo-mpo.gc.ca](mailto:Divya.Varkey@dfo-mpo.gc.ca) (D. Varkey), [Paul.Regular@dfo-mpo.gc.ca](mailto:Paul.Regular@dfo-mpo.gc.ca) (P. Regular), [Danny.Ings@dfo-mpo.gc.ca](mailto:Danny.Ings@dfo-mpo.gc.ca) (D. Ings), [Joanna.Flemming@dal.ca](mailto:Joanna.Flemming@dal.ca) (J. Mills Flemming).

<https://doi.org/10.1016/j.fishres.2021.105956>

Received 13 August 2020; Received in revised form 19 March 2021; Accepted 21 March 2021

Available online 4 April 2021

0165-7836/© 2021 Elsevier B.V. All rights reserved.

**Table 1**

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

| (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 49 cm 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 |

suggested the possibility of generating smooth ALKs using a Continuation Ratio Logit (CRL) model. Smooth ALKs have been applied before to lesser sandeel, (Rindorf and Lewy, 2001) North Sea haddock, (Stari et al., 2010; Berg and Kristensen, 2012) cod, herring and whiting (Berg and Kristensen, 2012). They can also mitigate the effects of the sampling artifacts mentioned above, demonstrated in Table 1b where the probability of being a particular age given a length of 49 cm is more suitably spread across nearby age classes and not just associated with age nine.

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

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

The spatial ALK approach presented here addresses the problem of smoothing over landmasses by using an approximation of a Gaussian Random Field (GF) that has support for physical barriers (Bakka et al., 2019). Desirably it still allows ALKs to be constructed at any location within the study area. Approximations of GFs have previously been proposed to help model spatial indices of abundance (Thorson et al., 2015; Thorson and Barnett, 2017), species distribution (Bakka et al., 2016) along with other non-marine uses such as global temperature data (Lindgren and Rue, 2011).

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

## 2. Methods

### 2.1. ordinal regression and continuation ratio logits

An ALK can be constructed using any classifier capable of handling multiple age classes and returning probabilities of a fish with a particular set of covariates (e.g. male, caught using gillnet, etc.) belonging to each age class. The ALK is simply those probabilities for every set of observed covariates. Ages naturally have an ordering associated with them, a seven year old fish must have first been a six year old fish and before that a five year old fish and so on. Ordinal regression can handle ordered categorical data and return class probabilities (Agresti, 2003). It is also possible to incorporate spatial structure directly into ordinal regression models through the use of splines or random fields. This is not

the case for some alternative multi-class classifiers like Classification and Regression Trees (CARTs). A number of different ordinal regression methods exist such as cumulative logits, adjacent-category logits, etc. Continuation Ratio Logits (CRLs) are one method that offers a few advantages (Agresti, 2003) (Harrell, 2014, pp. 311–312, 321–322).

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

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

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

where

$$\pi_a[\mathbf{x}_i] = \frac{p_a[\mathbf{x}_i]}{p_a[\mathbf{x}_i] + \dots + p_A[\mathbf{x}_i]} \quad (3)$$

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

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

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

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

### 2.2. Random fields

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

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

where  $\Gamma$  is the Gamma function,  $\nu$  is a smoothness parameter,  $K_\nu$  is the modified Bessel function of the second kind with order  $\nu$ ,  $r$  is the range parameter which is the spatial distance when the correlation is  $\approx 0.13$ ,  $\sigma_u$  is the marginal standard deviation (Bakka et al., 2018) and  $\|\mathbf{s} - \mathbf{t}\|$  is the distance between two points.

Using a GF directly is not computationally tractable for large problems. The cost to factorize the resulting dense covariance matrix is cubic in time. As a result a number of alternative approaches to try and get around the high computational cost have been proposed. Gaussian

Markov Random Fields (GMRFs) are GFs with the Markov property, that is

$$P(s_i | \{s_j : j \neq i\}) = P(s_i | x_j : j \in \mathcal{N}_i) \tag{6}$$

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

### 2.2.1. Gaussian random field approximation using stochastic partial differential equations

Lindgren and Rue (2011) found an explicit link between GFs and GMRFs when using a Matérn covariance function. A valid positive semi-definite covariance matrix is the result of the solution of a set of Stochastic Partial Differential Equations (SPDEs), which creates an approximation of a GF using a GMRF. This allows the benefits of modelling as a GF with the computational speed of a GMRF.

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

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

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

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

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

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

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

where  $\Omega_n$  is the set of nodes outside the boundary,  $\Omega_b$  is the set of nodes inside the boundary, and  $r_b$  is not a new range parameter but instead a predetermined fraction of  $r$ . In this case  $r_b = \frac{1}{10}r$ . This has the effect of essentially making the decorrelation range close to zero for points that

fall within the barrier creating the desired boundary properties. Other parameters are the same as described above. Further details on solving the SPDE for the GF approximation can be found in Lindgren and Rue (2011), Bakka et al. (2019) and Bakka (2018).

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

### 2.3. GF spatial age-length key

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

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

where  $\alpha_a$  is the intercept for age  $a$ ,  $\beta_a$  is the length parameter for age  $a$  and  $\xi_{a,s}$  is the spatial intercept resulting from the GF at location  $s$ :

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

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

### 2.4. Estimation

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

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

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

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

The optimal choice of  $\lambda$  can be chosen by k-fold cross validation (CV). This can be quite time intensive for large spatial models. Due to the presence of random effects in the spatial ALK model, an approach like

Generalized Cross Validation (GCV) which would avoid the time consuming k-fold CV can also not be used due to the difficulty in finding an influence matrix if it even exists at all. Instead a modified Akaike Information Criterion (AIC) is proposed here in place of k-fold CV. The modified AIC used here is defined to be

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

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

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

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

### 2.5. Simulation study

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

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

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

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

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

a GAM based model similar to the one presented in Berg and Kristensen (2012),

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

where  $f$  is a function of location  $s$  using thin-plate regression splines and finally the model proposed here as described in Eq. (10). An automatic selection of the maximum basis dimension ( $k$ ) is used for the thin-plate regression splines with AIC smoothness selection. The automatic selection is based on the method used by the DATRAS package discussed in Berg and Kristensen (2012) where the maximum basis dimension is the number of unique observations of the covariates appearing in the smooth terms minus one (i.e. the number of unique (non-zero) tow locations minus one) unless there are less than 10 unique locations for which then it falls back to a GLM. If including spatial information increases the accuracy of predicting what age class a fish belongs to, then the stratified survey estimates of abundance at age should also be closer to true abundance numbers at age. For all of the approaches age 10 was taken to be a plus group.

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

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

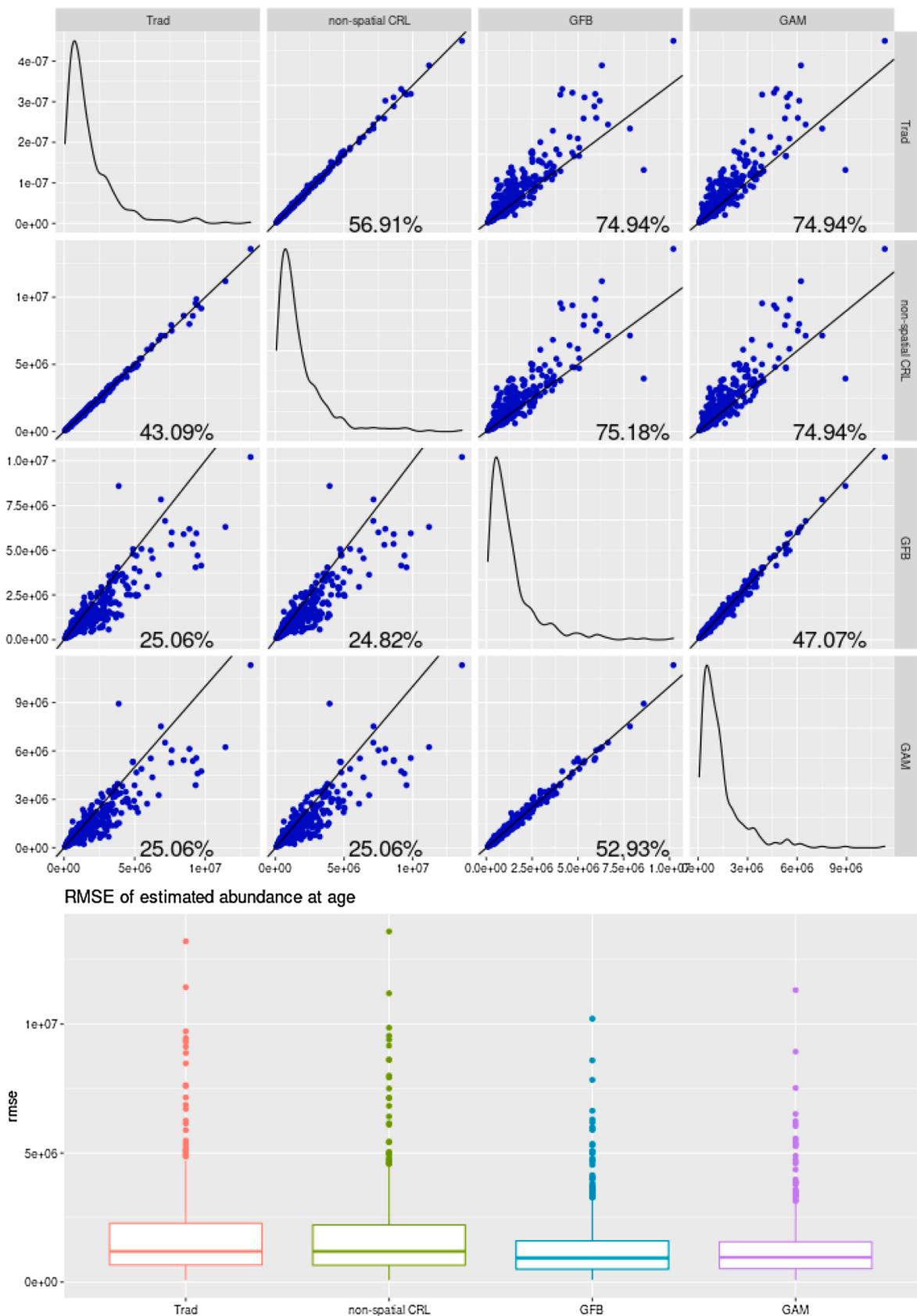
For each simulation the total stratified abundance at age is calculated. This follows the methods outlined in Smith and Somerton (1981) that developed from stratified random sampling techniques described in greater detail in books like Cochran (1977) and Lohr (2009). The survey area is divided into 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

$$\bar{Y}_{ah} = \frac{\sum_{i=1}^{N_h} y_{ahi}}{N_h} \tag{17}$$

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

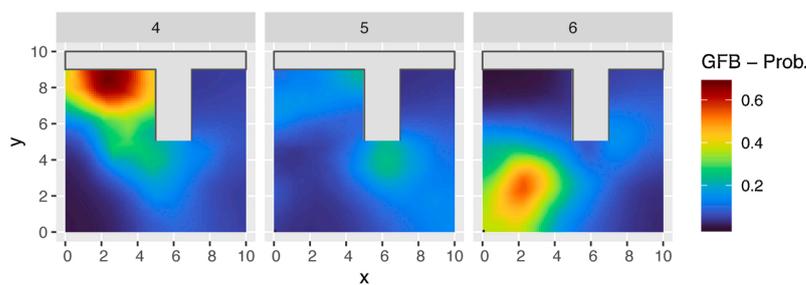
$$\hat{Y}_a = N \sum_{h=1}^H \frac{N_h}{N} \bar{Y}_{ah} \tag{18}$$

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

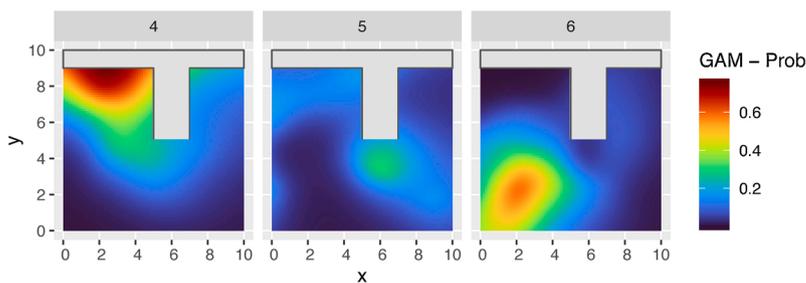


**Fig. 1.** The RMSE of the stratified survey estimate of abundance at age versus the true total abundance at age in the simulation. The top figure compares the RMSE of each of the methods against one another with a diagonal line illustrates where points would lie if the two methods were identical. The 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.

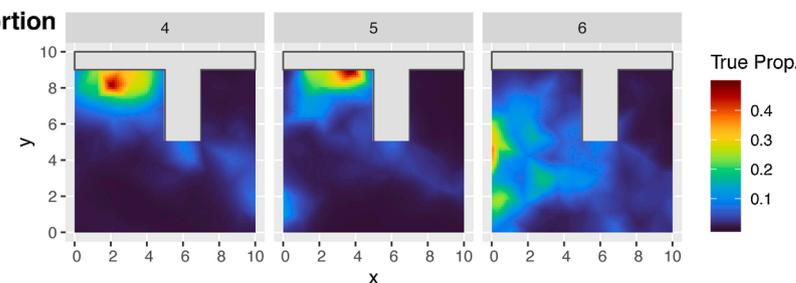
**GFB Prob.**



**GAM Prob.**



**True Proportion**



**Fig. 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.

**2.6. Application**

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

The mesh design can have an impact on the performance of the model. If a mesh is not dense enough, the approximation may not work well. Mesh designs can sometimes also impact the convergence of the model. Care should also be taken to ensure that all points that should be

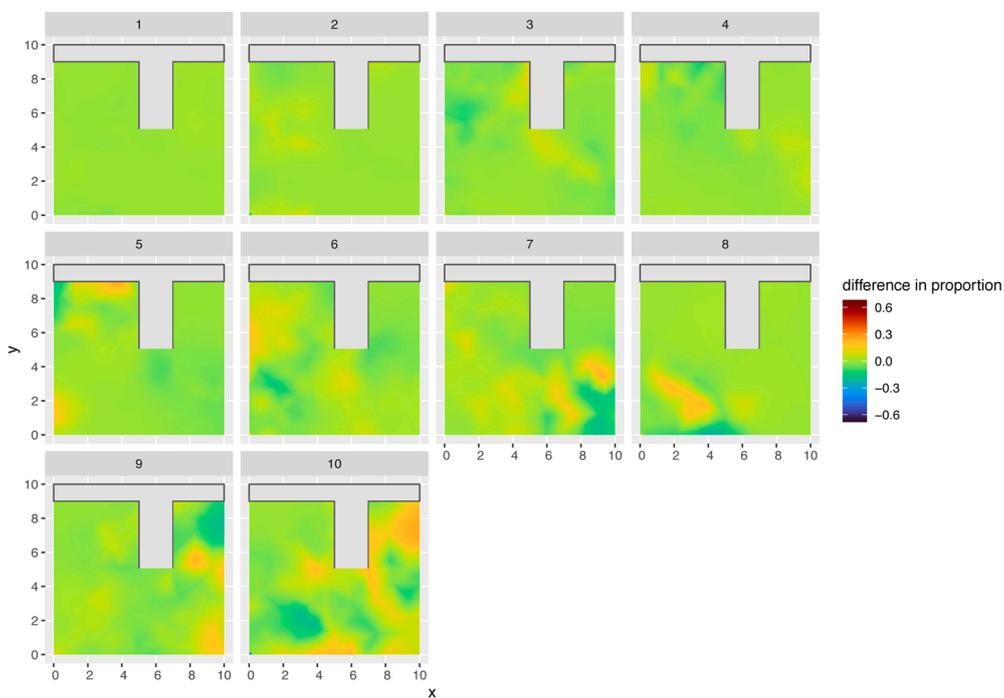
outside the boundary, are in fact. Further details on how to create a mesh for a barrier model can be seen in the referenced tutorial (Bakka, n. d.). As with the simulation in the previous section, age frequencies were generated using the same four methods and then stratified estimates of abundance were obtained.

**3. Results**

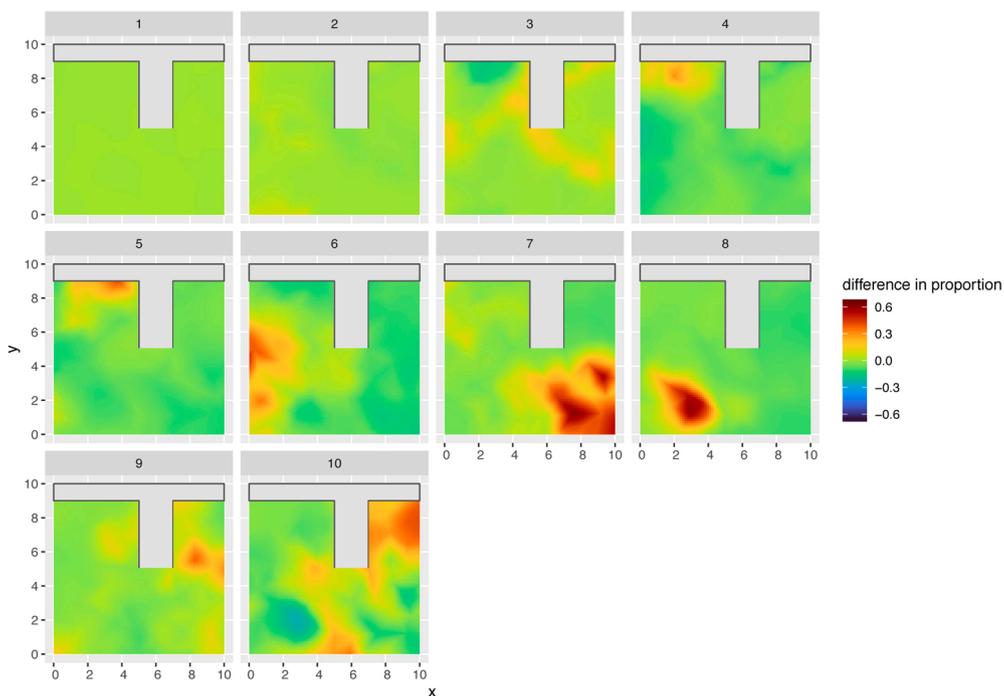
**3.1. Simulation study**

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

The bottom row of Fig. 2 is the true simulated spatial distribution for fish aged 4,5 & 6 for one simulation. The ages are somewhat overlapping in their spatial distribution but are centred in different areas. The spatial ALKs are better able to discriminate between ages by considering the location of the samples. This is evident in the top row of Fig. 2 which shows the probability of being a fish being age 4, 5 or 6 given a length of 40 cm across the simulated study area as obtained by the GFB model. The probability of being in that age class increases when predicting over the main bulk of that age class. The overall trend of the GAM approach is very similar to that of the GFB model while also displaying evidence of undesired smoothing underneath landmasses. Both the GFB and GAM yield a higher probability for 5 year olds south of the peninsula than

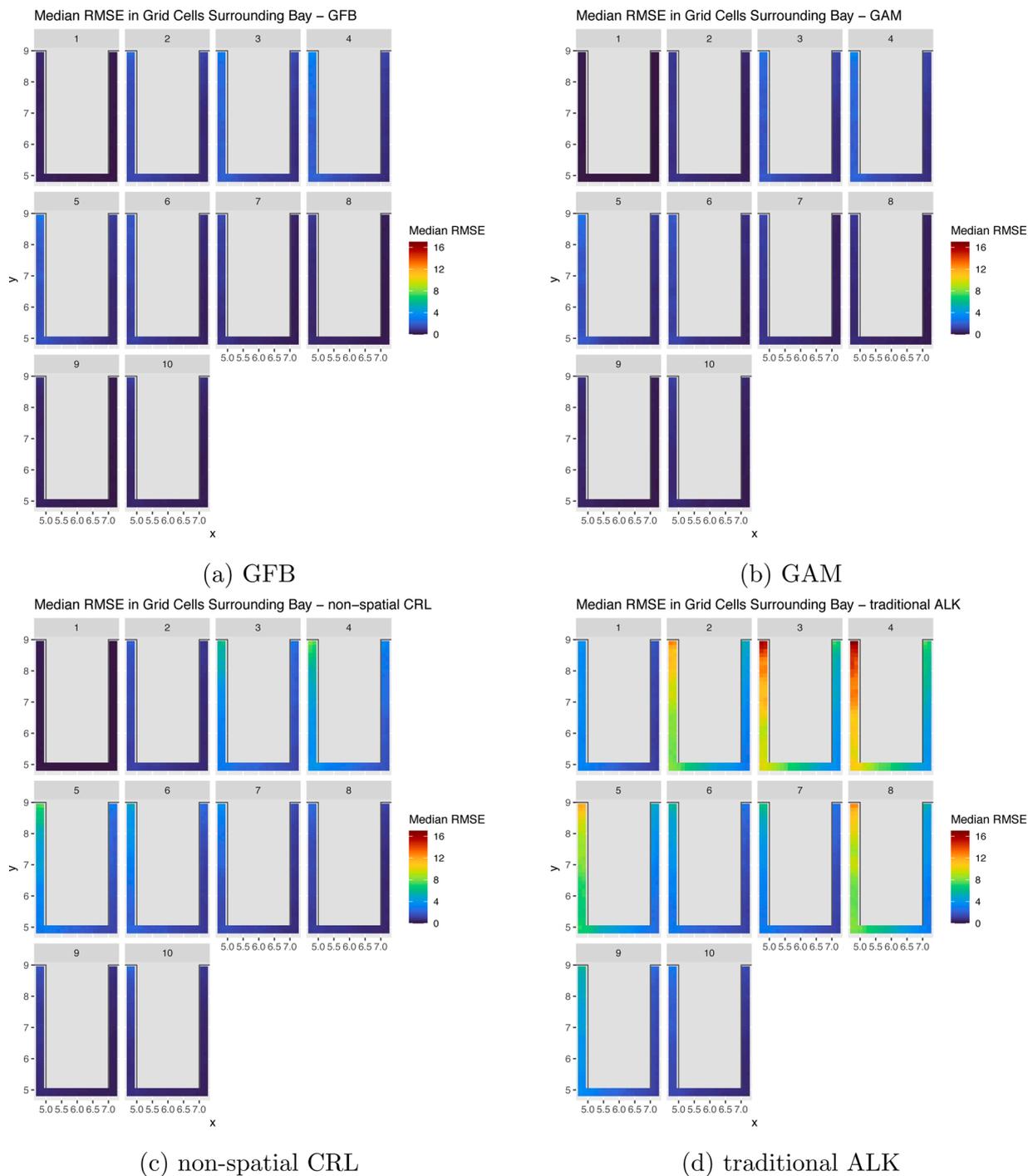


(a) Estimation error (true minus predicted) for the GFB model. A new spatial ALK is generated and applied to each simulation grid cell.



(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.

**Fig. 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 Fig. 2 for fish available to the survey. A perfect model would be a flat green with no difference in proportion between the two.



**Fig. 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.

**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 |

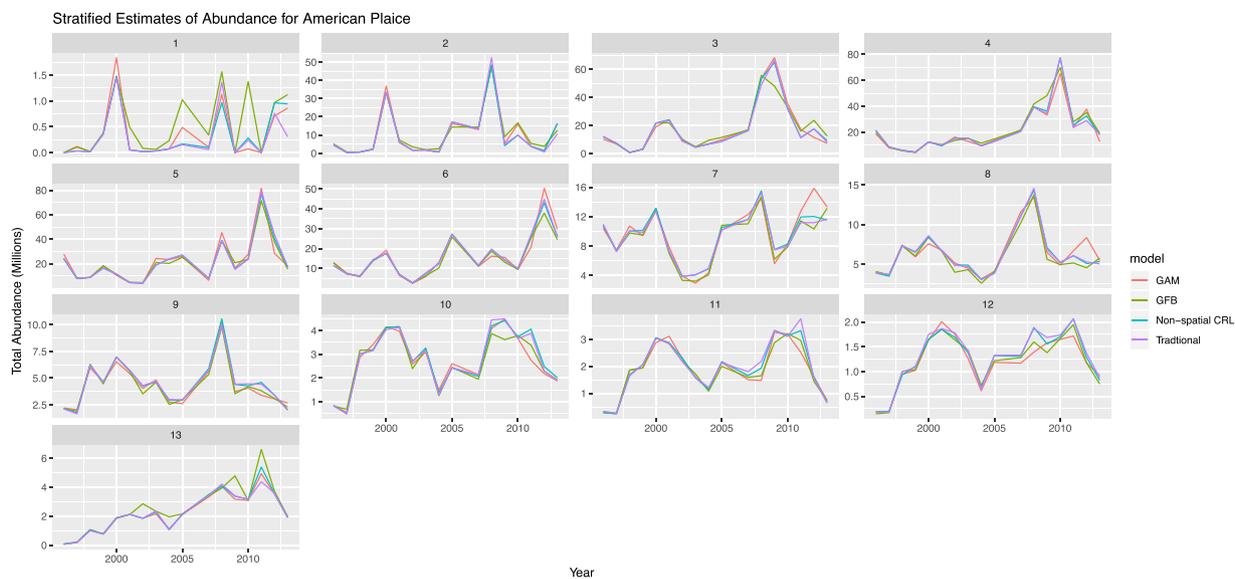


Fig. 5. Total abundance estimates by age for American Plaice within NAFO division 3P for the years from 1996 to 2013, not including 2006. Abundance estimates at age are largely similar across all four methods and follow the same general trends.

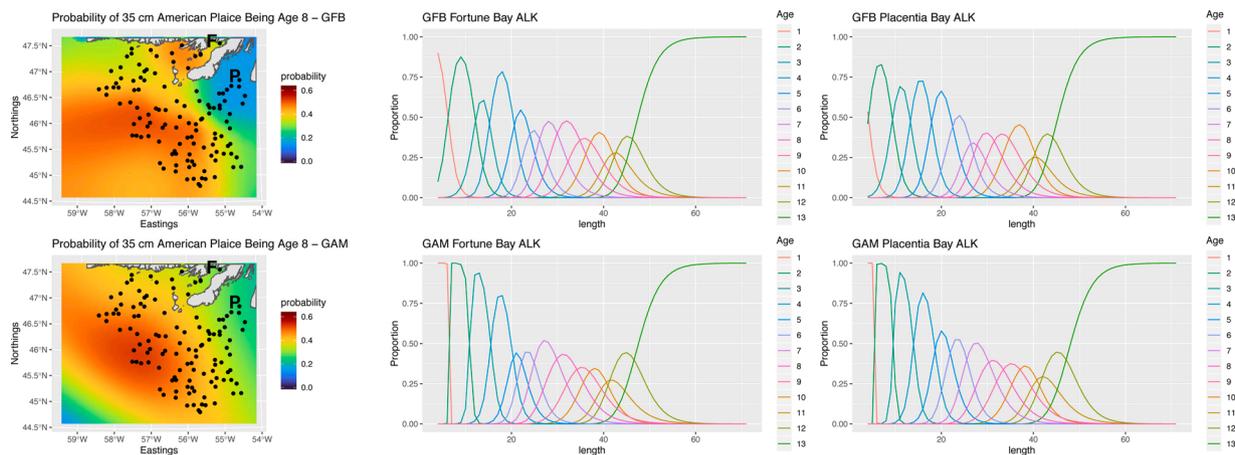


Fig. 6. A visual representation of three ALKs generated using the GF & GAM methods at the two locations shown (P-Placentia Bay, F-Fortune Bay) for the 2003 survey year. Each vertical slice of the graphs in columns 2 and 3 must sum to 1 and is like a row in an ALK. The maps are also surfaces showing the 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.

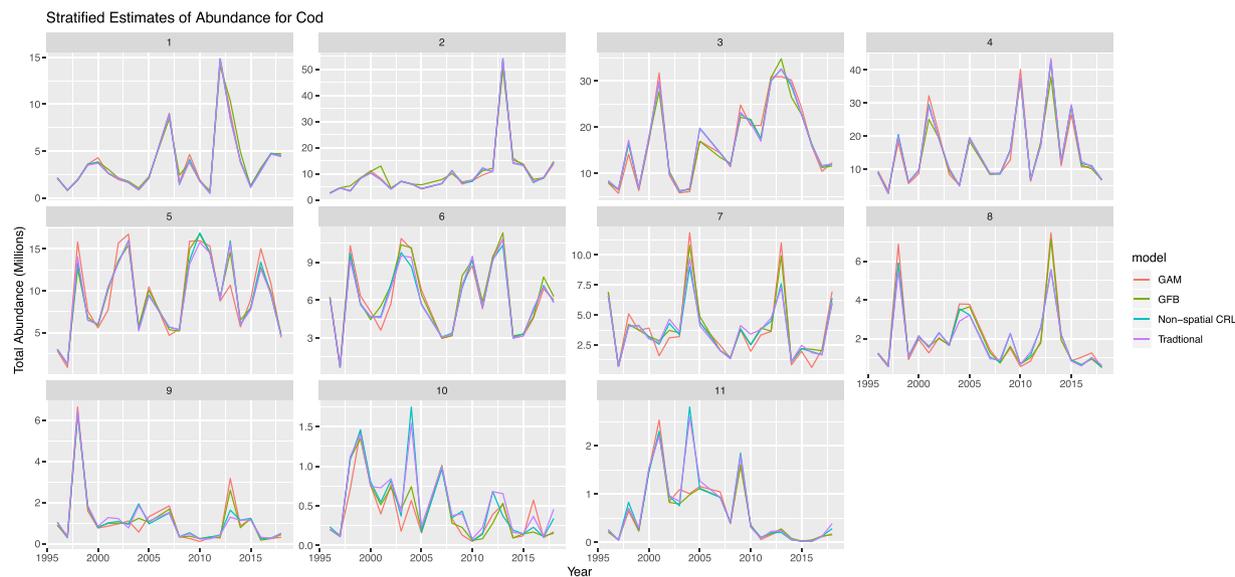
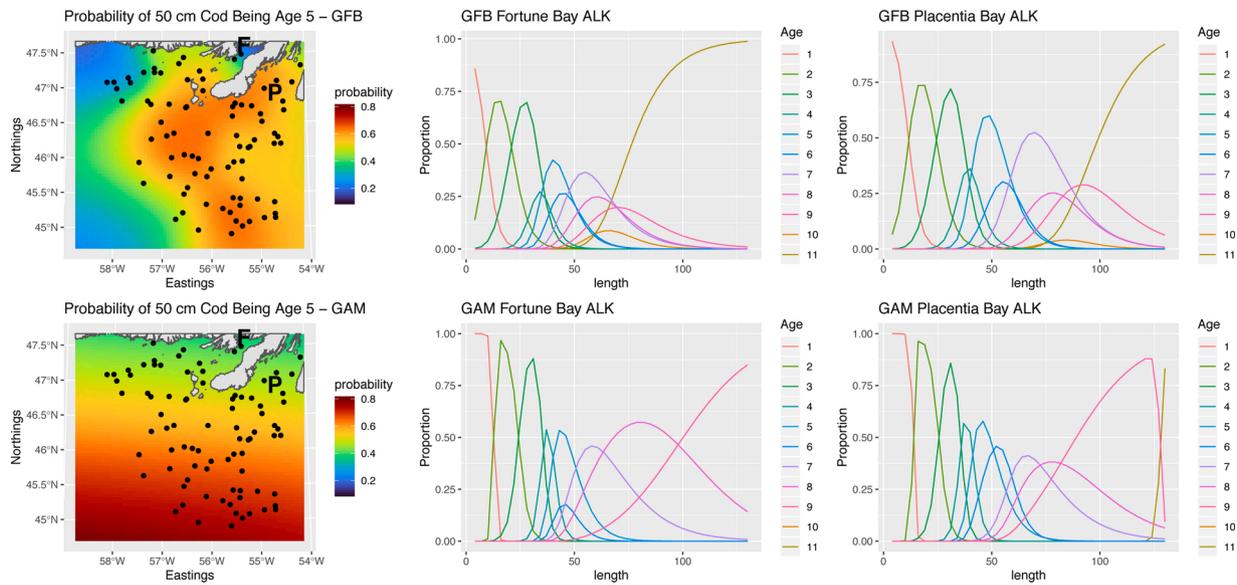


Fig. 7. Total abundance estimates by age for the cod dataset for the years from 1996 to 2018, not including 2006. Trends across years are broadly similar between methods.



**Fig. 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.

might be expected by the true proportions. This is due to the fact that the survey caught a larger share of age 5s than for other age classes.

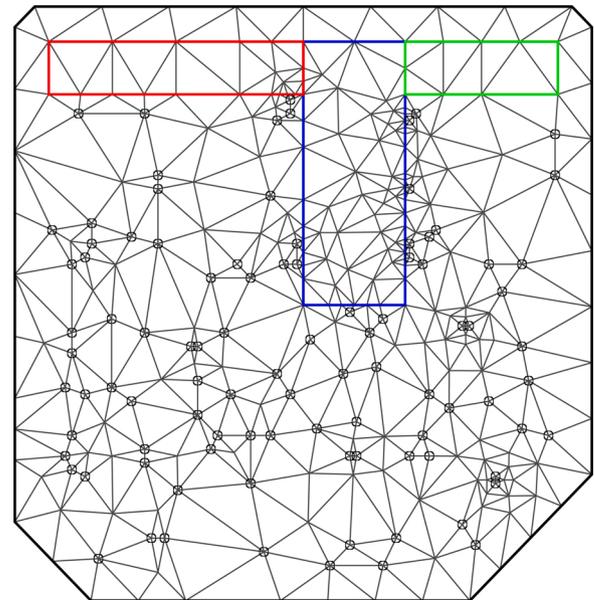
Each fish’s length in the population is simulated when the population is generated and then distributed spatially into the simulation grid cells such that the age and length for every fish in the simulation is known. The true proportion at age in each of the 4833 simulation grid cells was compared to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

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

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

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

### Constrained refined Delaunay triangulation



**Fig. 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.

### 3.2. Application

Atlantic cod and American Plaice are distributed throughout NAFO subdivision 3Ps, but during most years abundances is highest at particular locations such as the Halibut Channel (Cod) or the southeast slope of St. Pierre Bank (Plaice). The ALK methods discussed above were applied to both Cod and American Plaice data from DFO’s multi-species bottom trawl RV survey of NAFO subdivision 3Ps (Ings et al., 2019; Morgan et al., 2020). For the model based methods, the same model

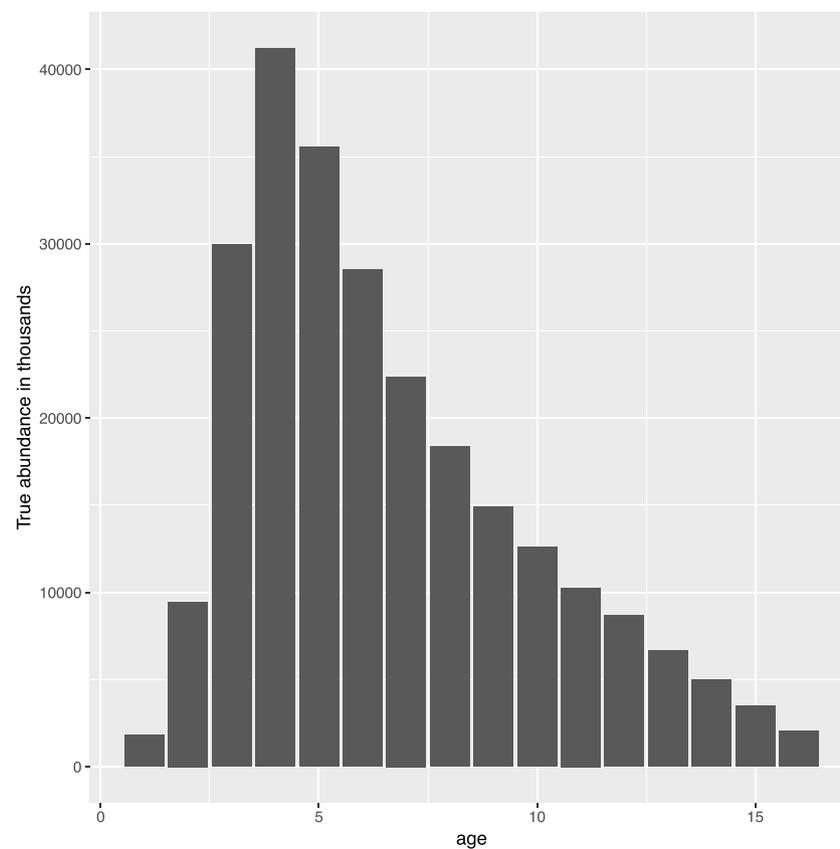


Fig. 10. The true abundance at age (in thousands) for the simulation used in Figs. 3 and 2.

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

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

### 3.2.1. American Plaice

American Plaice are associated with fine substrates and both juveniles and adults frequently occur in the same habitats (Morgan, 2000; Johnson, 2004). They do not conduct extensive annual migrations

(Johnson, 2004). The model based methods were fit from ages one to thirteen except for a handful of years where no age one Plaice otoliths sampled. In those cases the models were run on ages two through thirteen. Estimates of total abundance at age for American Plaice for the four different methods are shown in Fig. 5. With the exception of age one plaice, the four methods result in very similar estimates for the majority of the time series of total abundance at age (obtained by aggregating across space using the Stratified Mean Method (SMM)). Closer to the end of the time series there is a divergence for some age groups like 7 and 8 due to data sparsity.

Despite the similarity in aggregated abundance metrics across methods, when looking at the spatial GAM and GFB results there are differences in the predicted probabilities. For example in Fig. 6, the 2003 survey year the unconditional probabilities of an American Plaice being age 8 with a length of 35 cm were predicted across the space for both the GFB and GAM models. There does appear to be evidence of the probabilities being smoothed underneath the peninsula and into the bay in the GAM version that does not occur with the GFB models due to its support for physical barriers. Both models make it clear that the probabilities are spatially varying and there is a dependence on location. When looking at other ages not shown here, the two spatial methods do not always agree, the GAM method will sometimes predict almost flat gradients across space while the GFB for the same year will vary more across space. Fig. 6 also showcases examples of spatial ALK constructed from the GFB and GAM models at two different points. One at the tip of Placentia Bay, and one in Fortune Bay. Each curve represents the proportion at each length taken up by that age (like the columns of an ALK) while each vertical slice at a length must sum to one (like the rows of an ALK). Since age 13 was used as a plus group as fish get longer they end up more likely to land there. The two ALKs are different at each of the two points. For instance, the ALKs estimated by the GFB model shows more overlap between ages 1 and 2 in Placentia Bay than Fortune Bay while still having more overlap between the first two ages when compared against

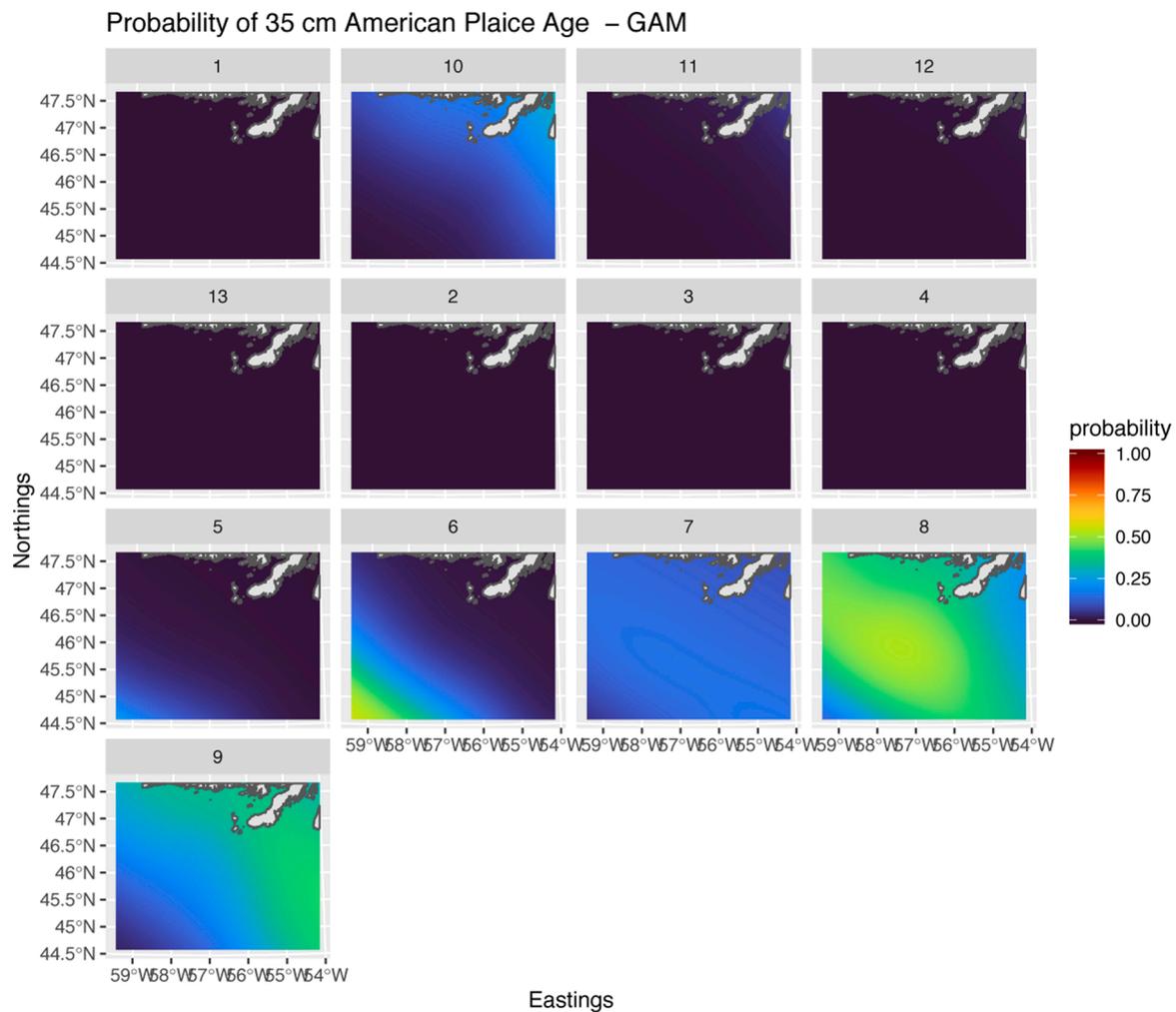


Fig. 11. The Probability of an American Plaice being each age class in the study area given a length of 35 cm as predicted by the GAM model.

the ALKs at the same points from the GAM model.

### 3.2.2. Cod

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

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

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

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

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

## 4. Discussion

Errors related to the process of obtaining age estimates are often ignored in age structured stock assessment models. Model based approaches offer an avenue for incorporating such errors even when a non-spatial method of aging is used. If errors in the aging process are ignored estimates obtained from the stock assessment model provide a false sense of precision and may impact derived quantities (e.g. spawning stock biomass). The simulation study here also suggests that spatial

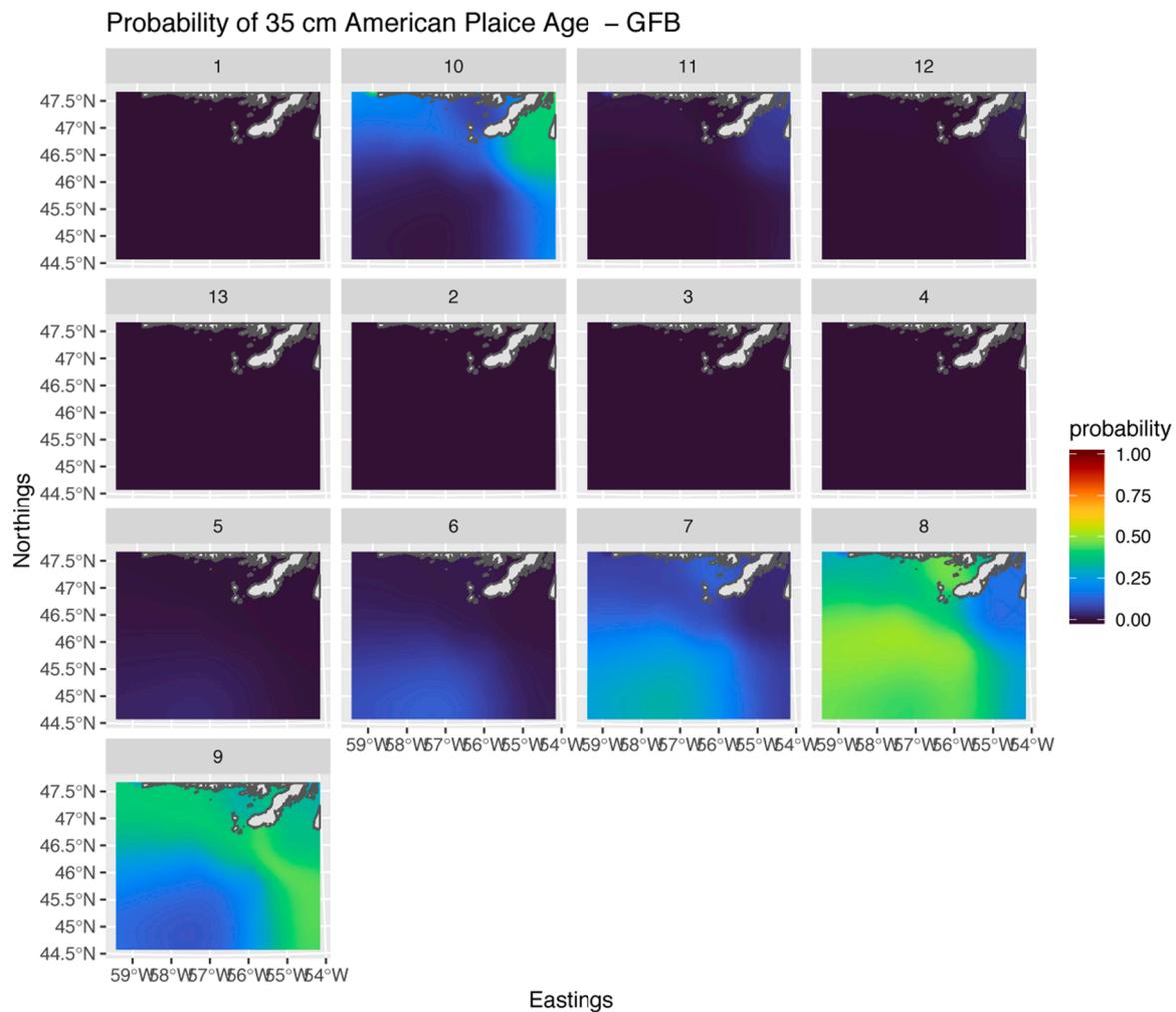


Fig. 12. The Probability of an American Plaice being each age class in the study area given a length of 35 cm as predicted by the GFB model.

methods have the potential to further reduce errors resulting from applying an ALK. Future work includes incorporating a spatial ALK model into a stock assessment model directly to see exactly how derived quantities and associated errors may be affected by the age estimation process.

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

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

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

The GFB had the lowest RMSE more often than the other three

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

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

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

**Credit statement**

**Jonathan Babyn:** Conceptualization, Methodology, Software, Formal Analysis, Writing – Original Draft.

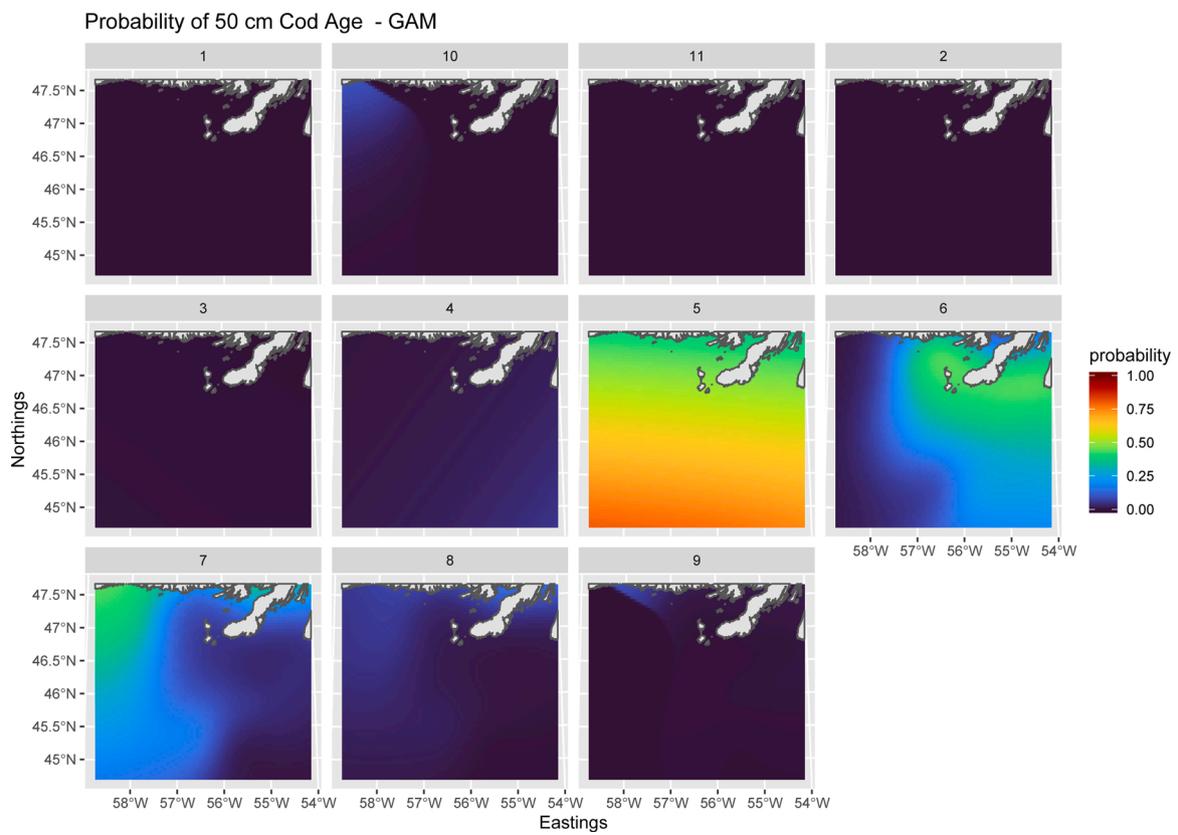


Fig. 13. The Probability of a Cod being each age class in the study area given a length of 50 cm as predicted by the GAM model.

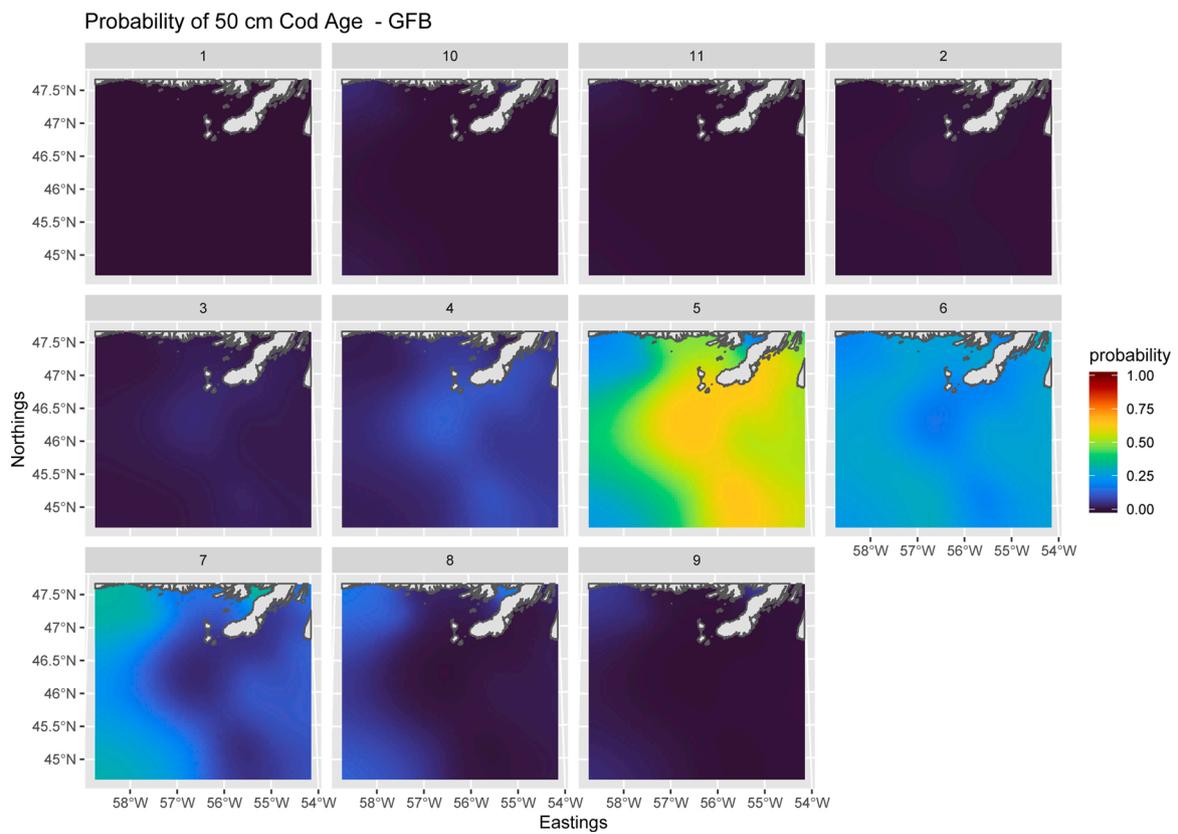
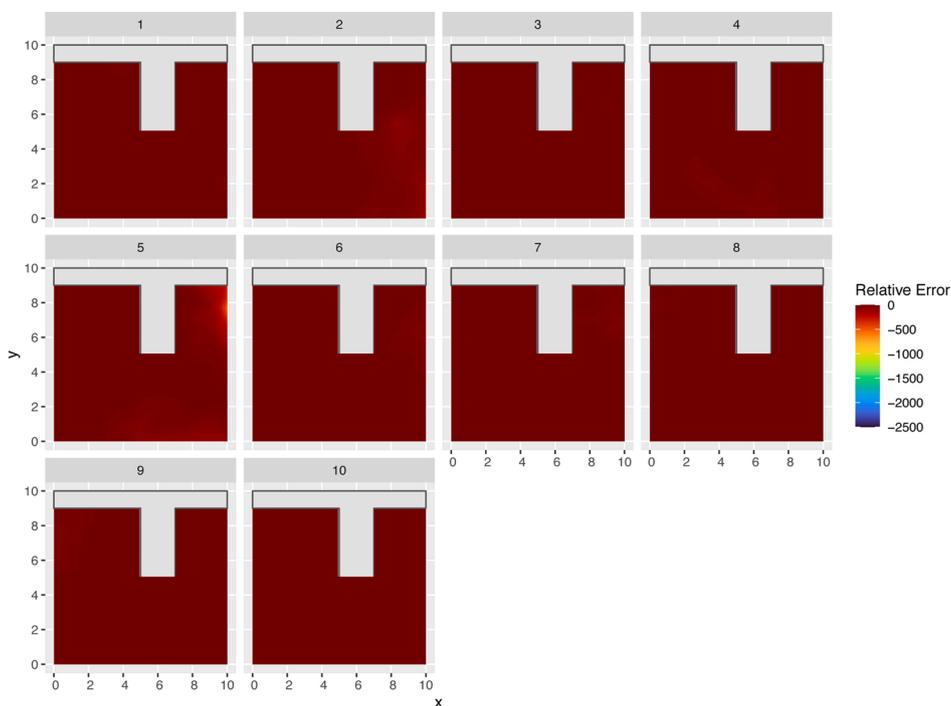
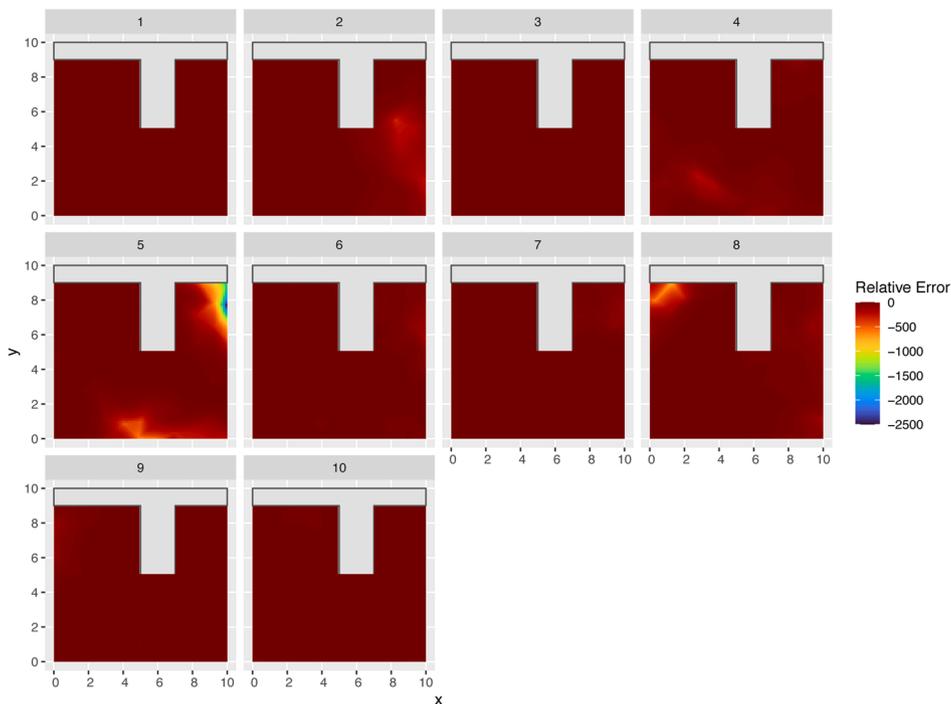


Fig. 14. The Probability of a Cod being each age class in the study area given a length of 50 cm as predicted by the GFB model.



(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.

**Fig. 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.

**Divya Varkey:** Conceptualization, Writing – Review & Editing.

**Paul Regular:** Conceptualization, Software, Writing – Review & Editing.

**Danny Ings:** Conceptualization, Resources, Writing – Review & Editing.

**Joanna Mills Flemming:** Supervision, Conceptualization, Writing – Review & Editing.

### Declaration of Competing Interest

The authors report no declarations of interest.

### Appendix A. Simulation code

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

In addition the Age-year-space covariance discussed in Appendix S3 of Regular et al. (2020) was instead obtained using a GMRF approximation with support for physical barriers as described in Section 2.2.1. A precision matrix  $Q$  is generated from a mesh and a specified set of hyperparameters. This is then approximated for each grid point in the `SimSurvey` simulation. This has the benefit over the default `SimSurvey` method of constraining the simulation to also have to abide by any physical barriers present and also provides a speed boost by using a GMRF approximation when the mesh has less nodes than there are cells in the grid. An example of a mesh used in the simulation is shown in Fig. 9 and example of the true abundance at age for the simulation referred to in Figs. 3a and 2 is shown in Fig. 10 (Figs. 11–15).

### References

- Aanes, S., Vølstad, J.H., 2015. Efficient statistical estimators and sampling strategies for estimating the age composition of fish. *Can. J. Fish. Aquat. Sci.* 72 (6), 938–953. <https://doi.org/10.1139/cjfas-2014-0408>.
- Aeberhard, W.H., Mills Flemming, J., Nielsen, A., 2018. Review of state-space models for fisheries science. *Annu. Rev. Stat. Appl. Sci.* 5, 215–235. <https://doi.org/10.1146/annurev-statistics-031017-100427>.
- Agresti, A., 2003. *Categorical Data Analysis*, vol. 482. John Wiley & Sons. <https://doi.org/10.1002/0471249688>.
- Bakka, H., 2018. How to Solve the Stochastic Partial Differential Equation That Gives a Matérn Random Field Using the Finite Element Method.
- H. Bakka, Mesh Creation including Coastlines, [https://haakonbakkagit.github.io/btopi\\_c104.html](https://haakonbakkagit.github.io/btopi_c104.html).
- Bakka, H., Rue, H., et al., 2018. Spatial modelling with R-INLA: a review. *Wiley Interdiscip. Rev.: Comput. Stat.* 10 (6).
- Bakka, H., Vanhatalo, J., Illian, J., et al., 2016. Accounting for Physical Barriers in Species Distribution Modeling With Non-stationary Spatial Random Effects (arXiv preprint). arXiv:1608.03787.
- Bakka, H., Vanhatalo, J., Illian, J.B., et al., 2019. Non-stationary Gaussian models with physical barriers. *Spat. Stat.* <https://doi.org/10.1016/j.spasta.2019.01.002>.
- Berg, C.W., Kristensen, K., 2012. Spatial age-length key modelling using continuation ratio logits. *Fish. Res.* 129, 119–126. <https://doi.org/10.1016/j.fishres.2012.06.016>.
- Brattley, J., Porter, D., George, C., 2002. Exploitation Rates and Movements of Atlantic Cod (*Gadus morhua*) in NAFO Subdiv. 3Ps Based on Tagging Experiments Conducted During 1997.
- Cochran, W.G., 1977. *Sampling Techniques*. John Wiley & Sons Inc.
- Dalley, E., Anderson, J., 1997. Age-dependent distribution of demersal juvenile Atlantic cod (*Gadus morhua*) in inshore/offshore northeast Newfoundland. *Can. J. Fish. Aquat. Sci.* 54 (S1), 168–176. <https://doi.org/10.1139/f96-171>.
- Fahay, M.P., et al., 1999. *Essential Fish Habitat Source Document. Atlantic cod, Gadus morhua, Life History and Habitat Characteristics*.
- Fridriksson, A., 1934. On the calculation of age-distribution within a stock of cod by means of relatively few age-determinations as a key to measurements on a large scale, *Rapports Et Proces-Verbaux Des Reunions. Conseil International Pour l'Exploration De La Mer* 86, 1–5.
- Gray, R.J., 1992. Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *J. Am. Stat. Assoc.* 87 (420), 942–951. <https://doi.org/10.1080/01621459.1992.10476248>.
- Harrell, F.E., 2014. *Regression Modeling Strategies, As Implemented in R Package 'rms' Version 3(3)*. <https://doi.org/10.1007/978-1-4757-3462-1>.
- Hilborn, R., Ovando, D., 2014. Reflections on the success of traditional fisheries management. *ICES J. Mar. Sci.* 71 (5), 1040–1046. <https://doi.org/10.1093/icesjms/fsu034>.
- Ings, D., et al., 2019. Assessing the status of the cod (*Gadus morhua*) stock in NAFO Subdivision 3Ps in 2018. *Fisheries & Oceans Canada, Science. Canadian Science Advisory Secretariat*.
- Johnson, D., 2004. *American Plaice, Hippoglossoides Platessoides, Life History and Habitat Characteristics*. NOAA Tech. Mem. NMFS-NE 187.
- Kristensen, K., et al., 2016. TMB: Automatic differentiation and laplace approximation. *J. Stat. Softw.* 70 (5), 1–21. <https://doi.org/10.18637/jss.v070.i05>.
- Kvist, T., Gislason, H., Thyregod, P., 2000. Using continuation-ratio logits to analyze the variation of the age composition of fish catches. *J. Appl. Stat.* 27 (3), 303–319. <https://doi.org/10.1080/02664760021628>.
- Lindgren, F., Rue, H., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B.* <https://doi.org/10.1111/j.1467-9868.2011.00777.x>.
- Lohr, S.L., 2009. *Sampling: Design and Analysis*. Nelson Education. <https://doi.org/10.1201/9780429296284>.
- Morgan, M.J., et al., 2020. Assessing the status of the American Plaice (*Hippoglossoides platessoides*) stock in NAFO Subdivision 3Ps in 2019. *Fisheries & Oceans Canada, Science. Canadian Science Advisory Secretariat*.
- Morgan, M., 2000. Interactions between substrate and temperature preference in adult American plaice (*Hippoglossoides platessoides*). *Mar. Freshw. Behav. Phys.* 33 (4), 249–259. <https://doi.org/10.1080/10236240009387096>.
- Parrish, J.K., 1999. Using behavior and ecology to exploit schooling fishes. *Environ. Biol. Fish.* 55 (1–2), 157–181.
- Punt, A.E., 2019. Modelling recruitment in a spatial context: a review of current approaches, simulation evaluation of options, and suggestions for best practices. *Fish. Res.* 217, 140–155. <https://doi.org/10.1016/j.fishres.2017.08.021>.
- Punt, A.E., Haddon, M., Tuck, G.N., 2015. Which assessment configurations perform best in the face of spatial heterogeneity in fishing mortality, growth and recruitment? A case study based on pink ling in Australia. *Fish. Res.* 168, 85–99. <https://doi.org/10.1016/j.fishres.2015.04.002>.
- Regular, P.M., et al., 2020. `SimSurvey`: an R package for comparing the design and analysis of surveys by simulating spatially-correlated populations. *PLOS ONE* 15 (5), e0232822. <https://doi.org/10.1371/journal.pone.0232822>.
- Rindorf, A., Lewy, P., 2001. Analyses of length and age distributions using continuation-ratio logits. *Can. J. Fish. Aquat. Sci.* 58 (6), 1141–1152. <https://doi.org/10.1139/f01-062>.
- Ross, S.M., 2014. *Introduction to Probability Models*. Academic press.
- Rue, H., Held, L., 2005. *Gaussian Markov Random Fields: Theory and Applications*. CRC Press. <https://doi.org/10.1201/9780203492024>.
- Smith, S., Somerton, G., 1981. *STRAP: A User-Oriented Computer Analysis System for Groundfish Research Trawl Survey Data*, Technical Report. Department of Fisheries and Oceans.
- Stari, T., et al., 2010. Smooth age length keys: observations and implications for data collection on North Sea haddock. *Fish. Res.* 105 (1), 2–12. <https://doi.org/10.1016/j.fishres.2010.02.004>.
- Thorson, J.T., Barnett, L.A., 2017. Comparing estimates of abundance trends and distribution shifts using single-and multispecies models of fishes and biogenic habitat. *ICES J. Mar. Sci.* 74 (5), 1311–1321. <https://doi.org/10.1093/icesjms/fsw193>.
- Thorson, J.T., Shelton, A.O., et al., 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *ICES J. Mar. Sci.* 72 (5), 1297–1310. <https://doi.org/10.1093/icesjms/fsu243>.
- Verweij, P.J., Van Houwelingen, H.C., 1994. Penalized likelihood in Cox regression. *Stat. Med.* 13 (23–24), 2427–2436. <https://doi.org/10.1002/sim.4780132307>.
- Wood, S.N., Bravington, M.V., Hedley, S.L., 2008. Soap film smoothing. *J. R. Stat. Soc.: Ser. B (Stat. Methodol.)* 70 (5), 931–955. <https://doi.org/10.1111/j.1467-9868.2008.00665.x>.
- Worm, B., et al., 2009. Rebuilding global fisheries. *Science* 325 (5940), 578–585. <https://doi.org/10.1126/science.1173146>.