Variance Estimators of North Sea International Bottom Trawl Survey Indices

Abstract

1 INTRODUCTION

Fish stock assessments are used by fishery managers for making management decisions regarding catch quotas. The assessments provide fundamental information about the status of the stock, for instance, whether the stock is increasing and support for increased levels of harvest should be given, or whether the stock is decreasing and stricter control on harvest should be implemented. Associated with the parameters used in fish stock assessment is their uncertainty, which should not be ignored when formulating management policies (Walters and Ludwig, 1981; Ludwig and Walters, 1981). This uncertainty can arise from many sources including natural variability, estimation procedures and lack of knowledge regarding the parameter (Ehrhardt and Legault, 1997). The North Sea International Bottom Trawl Surveys (NS-IBTS) data, coordinated by the International Council for the Exploration of the Sea (ICES), provides information on seasonal distribution of stocks and estimates of abundance indices and catch in numbers of fish per age-class without an assessment of the accuracy of these estimates. As pointed out by Ludwig and Walters (1981) estimates of parameters relating to stock size are of little value unless they are accompanied by uncertainty estimates.

Indices of abundance at age from the NS-IBTS are based on data from a stratified cluster sampling approach, and it is essential to account for the sampling approach so as to produce reliable variance estimates (Lehtonen and Pahkinen, 2004). If the sampling approach is ignored, the effect on the variance of the parameters could be substantial. In particular, the variance could be greatly inflated due to the clustering effect, which involves intra-cluster correlation of the variables (Aanes and Vølstad, 2015; Lehtonen and Pahkinen, 2004). Currently, abundance indices from the NS-IBTS are estimated using an age-length key (ALK) method (Fridriksson, 1934), which is assumed to be constant over relatively large areas. In this paper we give a strong case for assuming variation in the ALK within these areas (see Figure 1, which shows the estimated age probabilities of a 30 cm cod (*Gadhus morhua*) in the first quarter of 2015). Figure 1 shows that the age distribution clearly varies for a 40 cm cod within Central North Sea and Northern North Sea. We propose two ALK estimators, which consider spatial variation: 1) a nonparametric ALK estimator, and 2) a spatial model-based ALK estimator, which we describe in Section 3. Section 1.1 gives an overview of the North Sea International Bottom Trawl Surveys. A brief description of the data is given in Section **??**. The current estimators for ALK and catch per unit effort (CPUE) used by ICES in their database for trawl surveys (DATRAS) and our proposed ALK estimators are given in Section 3. The results are given in Section 4 and a discussion is given in Section 5.

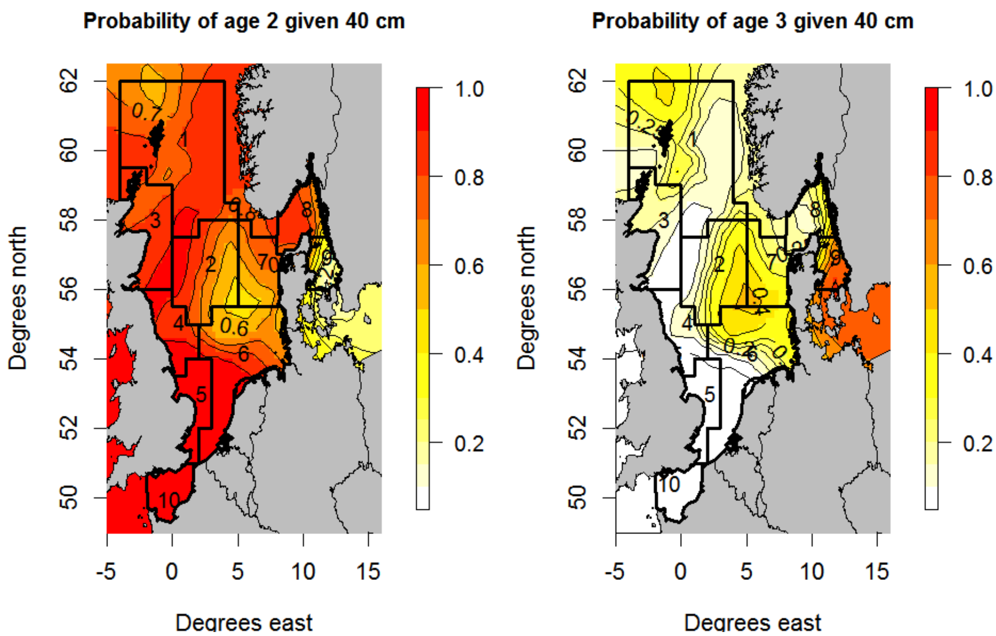

Figure 1: Estimated probability of age of a 0 cm long cod in the first quarter of year 201. The probability of age three or older is approximately zero. The polygones marked 1 to 10 is the round fish areas (RFAs) where the ALK is assumed constant in the currently used estimators of the official CPUEs.

1.1 Overview of the North Sea International Bottom Trawl Surveys

The North Sea International Bottom Trawl Surveys (NS-IBTS) was formed in 1991, which is a combination of the International Young Herring Survey (IYHS) and eight national surveys in the North Sea, Skagerrak and Kattegat areas. These surveys began in the 1960's, and the 1970's and 1980's, respectively. The IYHS was developed with the aim of obtaining annual recruitment indices for the combined North Sea herring *Clupea harengus* stock (ICES 2012), but yielded valuable information on other fish species such as cod *Gadus morhua* and haddock *Melanogrammus aeglefinus*.

The NS-IBTS began with quarterly surveys providing information on seasonal distribution of stocks sampled, hydrography and the environment, which allows changes in fish stock to be monitored and abundance of all fish species (Table 1) to be determined. These quarterly surveys, however became difficult to sustain as countries experienced budget cuts making it impossible to maintain high levels of research vessel effort. As such, in 1997 countries carried out a survey only twice a year; a first quarter survey (January-February) and a third quarter survey (August-September). Table 1 gives the common names (scientific names in parentheses) of the target species that are sampled during the quarterly North Sea International Bottom Trawl Surveys. The common names of the species in parentheses will be used in the rest of paper.

Table 1: Species fished in the NS-IBTS from 1991-2017.

Standard Pelagic	Standard Roundfish	By-Catch Gadoid
Herring (Clupea harengus)	Cod (Gadus morhua)	Pollock (Pollachius)
Sprat (Sprattus sprattus)	Haddock (Melanogrammus aeglefinus)	Pouting (Trisopterus luscus)
Mackerel (Scomber scombrus)	Norway Pout (Trisopterus esmarkii)	Trisopterus minutus (Poor Cod)
	Saithe (Pollachius virens)	Blue Whiting (Micromesistius poutassou)
	Whiting (Merlangius merlangus)	Hake (Merluccius merluccius)
		Ling (Molva molva)
		Tusk (Brosme brosme)

Research vessels from seven nations in the first quarter (Q1) and six nations in the third quarter (Q3) are used for conducting surveys on all finfish species in the North Sea during January-February and July-August, respectively, in 1997-2017. The sampling frame is defined by the ICES index or roundfish areas (RFA) as shown in Figure 2 numbered 1 to 10, and which we refer to as superstrata (Nottestad et al., 2015; Fuller, 2011). These roundfish areas were substratified into small strata defined by non-overlapping statistical rectangles of roughly 30×30 nautical miles (1^o Longitude $\times$ 0.5^o Latitude), and were convenient to use for NS-IBTS as they were already being used for fisheries management purposes. Most statistical rectangles contain a number of possible tows that are deemed free of obstructions, and vessels are free to choose any position in the rectangles as long as the hauls are separated by at least 10 nautical miles within and between rectangles. In some rectangles, sampling may be further stratified due to significant changes in seabed depth which may, in turn, cause variations in the fish population. In particular, the NS-IBTS herring, saithe and sprat data are weighted by depth strata in the statistical rectangle (Table **??**). It is also a requirement that countries avoid clustering their stations between adjacent rectangles in order to reduce positive serial correlation, and thereby maximize survey precision. The latest major reallocation of rectangles occurred in

1991, but since then the survey has tried to keep at least one vessel in every subarea in which it had fished in the most recent years. Minor reallocation of rectangles between Norway, Scotland and Germany was done in 2013. Each rectangle was typically sampled twice by two different countries before 1997, but after that target coverage of two trawl hauls per rectangle per survey was introduced because of national financial constraints (ICES 2017). But in some rectangles in the Eastern English Channel, Southern North Sea and Central North Sea intensified sampling is carried out.

The trawl tow locations are selected using a semi-random approach with at least two primary sampling units (PSU) per stratum, where PSUs are standardized swept-area trawl hauls. Sampling locations for all countries are proposed in advance in order to increase the randomisation of sampling. These locations are based on a random selection on a random selection of valid tows with start and end position executed in the period 2000-2017. In the unusual event that no "clear" tow exists the cruise leader, who select the haul positions, may select to undertake a "blind" tow on unknown ground after checking the proposed trawl track for hazardous seadbed obstructions with acoustic methods.

The recommended standard trawling gear of the NS-IBTS is the mulitpurpose chalut à Grande Ouverture Verticale (GOV) trawl (ICES 2012), which has been used on all participating vessels since 1992, while different pelagic and bottom trawls suitable for fishing finfish species were used before 1992. Standardized trawling protocols were adopted with a towing speed of 4 knots but depending on vessel performance, tide and weather conditions the average towing speed can be at minimum 3.5 and maximum 4.5 knots. From 2000-2018 trawling is done during the daylight hours, which are considered 15 minutes before sunrise to 15 minutes after sunset (ICES 2012). After each trawl the total catch of the different species is weighed on board and biological parameters such as length for all fish species caught (to 0.1cm below for shellfish, to 0.5cm below for herring and sprat and to 1cm below for all other species) are collected. Where the numbers of individuals are too large for all of them to be measured to obtain the length distribution, a representative subsample of 100 fish is selected. Otoliths are collected on board from a small fraction of all the target species from all round fish areas (RFAs) (Figure 2) to retrieve age reading. Table 4 in Appendix.... gives the minimum sampling levels of otoliths for the target species.

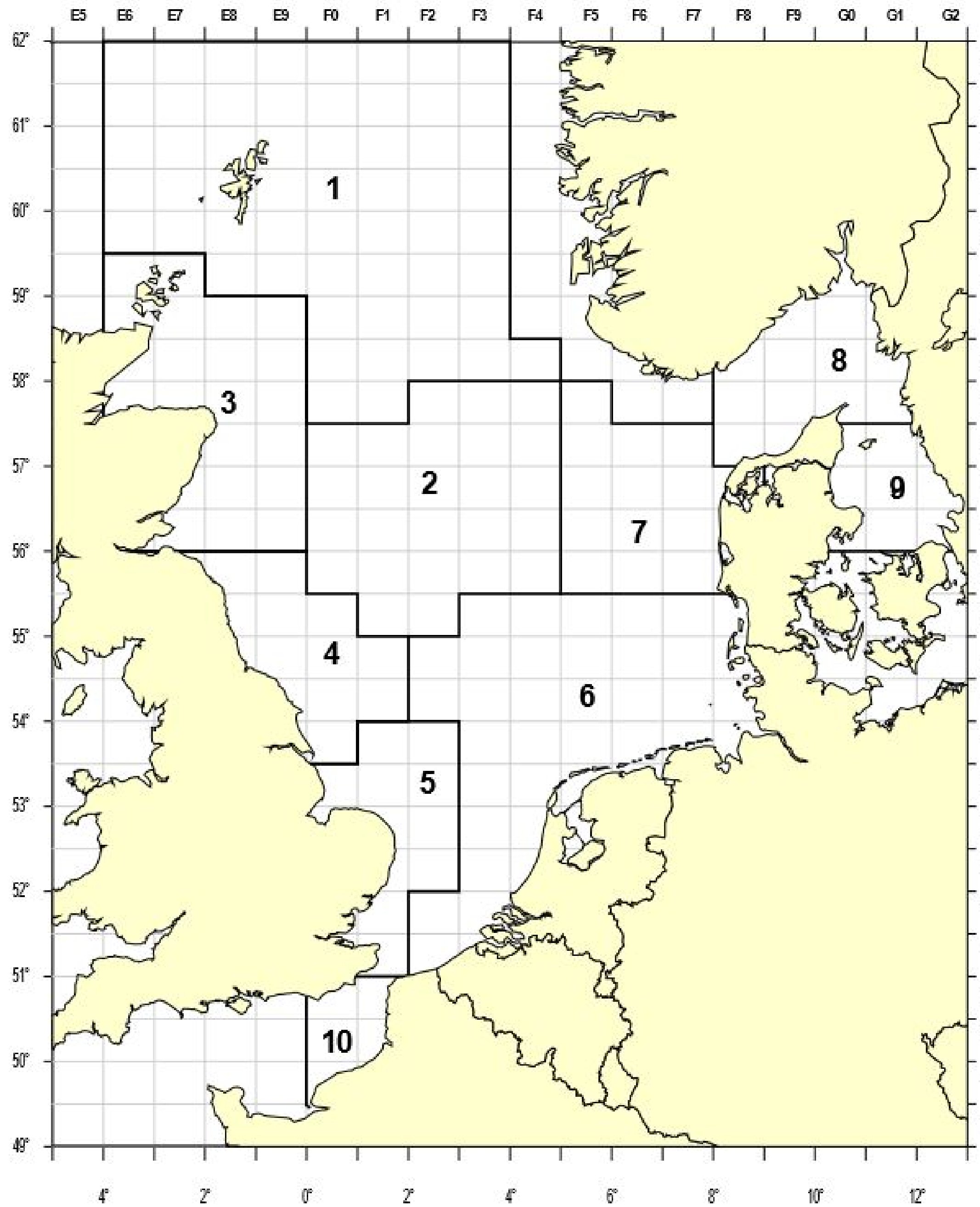

Figure 2: Standard roundfish areas used for roundfish since 1980, for all standard species since 1991. Additional RFA 10 added in 2009. For example, the number 1 indicates ICES Index Area 1, and an ICES Statitical rectangle (ST) in IA 1 is 43F1.

2 The North Sea Cod Data

An analysis of the North Sea Cod catches from the first quarter of IBST 2015 is presented. In general, the NS-IBTS data is registered as follows: 1) data calculated as catch in numbers per hour trawled (denoted as C type), 2) data by haul (denoted as R type), and 3) sub-sampled data (denoted as S type). For each species (Table 1) and by age group, abundance indices are calculated by averaging within statistical rectangles (strata) and then averaging over specific round fish areas (RFAs). Cod is typically found in all RFAs (see Figure 1). Table 2 gives an overview of the data used.

Table 2: Summary of NS-IBTS Cod data for first quarter of 2015.

Data	**Description**
Trawl hauls	Total of 387 trawl hauls (303 with age information of cod)
Age	The age of cod varied between 1 to 9 years.
Length	Length information in cm of each cod varied between 9 to 113 cm
Date	Date of catch varied between 13.01.2015 to 19.02.2015
Statistical rectangle	The stratum in which at least two trawl hauls are made
Coordinates	Geographic coordinates of each trawl haul in a statistical rectangle
Duration of haul	Mean duration is 25.9 minutes, with 15 to 30 minutes as 90% coverage interval.

3 METHODS

This section gives the estimators of abundance indices. The estimators are haul time-based and utlizes an ALK approach. We consider the ALK approach used in DATRAS and we propose two ALK estimators. The ALK used in DATRAS for computing abundance indices does not account explicitly for the spatial distribution in the age-length composition. To account for the spatial distribution we propose a design-based ALK estimator that is haul dependent (Section 3.2.2) and a model-based ALK estimator (3.2.3).

3.1 CPUE Estimators

For a given species of interest, define $n_{h,l}$ to be the number of fish with length l caught by the hth trawl haul. Define the CPUE for a given trawl h to be

$$\text{CPUE}_{h,l} = \frac{n_{h,l}}{d_h}, \tag{3.1}$$

were d_h is the duration of the trawl in hours. The mean CPUE in a statistical rectangle is further defined as the average of the CPUE for each trawl haul in the rectangle:

$$\text{mCPUE}_{s,l} = \sum_{h \in H_s} \frac{\text{CPUE}_{h,l}}{|H_s|}. \tag{3.2}$$

Here H_s represents the set of trawl hauls taken in statistical rectangle s, and $|H_s|$ is the number of hauls taken in the rectangle. The mean CPUE in pth RFA is further defined as

$$\text{mCPUE}_{p,l} = \sum_{s \in S_p} \frac{\text{mCPUE}_{s,l}}{|S_p|}, \tag{3.3}$$

where S_p is the set of all statistical rectangles in pth RFA and $|S_p|$ is the number of statistical rectangles in pth RFA.

The cpue per age class is further defined as

$$\text{CPUE}_{h,a} = \sum_{l \in \mathbf{L}} \text{CPUE}_{h,l} \times ALK_{a,l,h}, \tag{3.4}$$

where $ALK_{a,l,h}$ is an age length key which represents the estimated proportion of fish with age a in lth length class in haul h, and $\mathbf{L}$ is the set of all length classes. The mean CPUE per age in the statistical rectangles and roundfish areas are defined as (3.2) and (3.3) with $\text{CPUE}_{(\cdot),l}$ substituted by $\text{CPUE}_{(\cdot),a}$.

An estimator for catch at age in the North Sea is needed, and variance estimator

3.2 ALK Estimators

3.2.1 DATRAS ALK Estimator

Let ALK^{D} be the ALK used by DATRAS, which is currently used for producing offical CPUE per age estimates. The $ALK^{\text{D}}_{a,l,h}$ is defined as the proportion of observed fish with age a in length class l in the RFA. If there are no observed fish in length class l in the RFA, ages from length classes close to l is used. The details of the procedure for borrowing strength from neighbouring length classes is given in appendix C.1.

The underlying assumption of this ALK approach is that age-length compositions are homogeneous within the RFAs. This is a rather strong assumption, and any violation have an unknown impact on the estimates of abundance indices. In fact, Kimura (1977) showed that the application of an age-length key to a population where the age composition differs from that of the population from which the age-length key was drawn will give bias results. Because the ALK may be haul dependent we propose an ALK method that is based on trawl hauls, which we denote by ALK^H.

3.2.2 Haul Dependent ALK Estimator

We define a haul dependent ALK by ALK^H. The $ALK^H_{a,l^*,h}$ is defined as the average proportion of observed fish with age a in a pooled length class l^* in haul h. We use pooled length classes for this estimate since there are typically few observed length classes in a single haul. We define a pooled length class to consist of five length classes, the first pooled length class consist of the five smallest length classes and so on. If there are no observed ages of fish in a pooled length class l^* in the haul, ages from the same pooled length class in the haul closest in air distance from the hth haul is used. If there are no observed fish within the pooled length class in the closes haul, the next closes haul is used and so on. The details of borrowing strength from length classes in hauls closest in space is given in appendix C.2.

3.2.3 Spatial Model-Based ALK Estimator

The ALK approaches defined in Sections 3.2.1 and 3.2.2 use the method of "borrowing" age-length compositions within or between hauls for estimating abundance indices when data points for age-length combinations are missing. In this section we propose a statistical model-based approach to fill in missing values in an objective and robust manner, while accounting for the uncertainty that arises due to sampling variability (Berg and Kristensen, 2012). The statistical model allows the creation of a smooth distribution of age given length and location, and can include other covariates such as the random effect of the trawl haul. Spatial model-based approach of age-lengths has been widely used in fisheries assessment (Berg and Kristensen, 2012; Kvist et al., 2000; Rindorf and Lewy, 2001), where Continuous ratio logit (CRL) models were applied and where Generalized Linear Models (GLMs) have been used for estimation. We consider Logits (Dyke and Patterson, 1952; Agresti, 2003), which is a type of model for categorical response data (such as age groups) and, which have been previously used for modelling ALKs and estimating uncertainty (Gerritsen et al., 2006).

Let the response variable of the age group of a fish be $a = M, ..., A$ where M is the youngest age and

A is the oldest age, which is typically defined as a "plus group". Suppose $y(l, \mathbf{s}, h)$ is the age of a fish with length l, caught at location $\mathbf{s}$ by trawl haul h, then the the probability of age a in a given year and quarter is given by:

$$\pi_a(y(l, \mathbf{s}, h)) = \begin{cases} \frac{\exp(\mu_a)}{1+\sum_{i=M}^{A-1} \exp(\mu_a)}, & a < A \\ \\ \frac{1}{1+\sum_{i=M}^{A-1} \exp(\mu_a)}, & a = A. \end{cases}, \tag{3.5}$$

where

$$\mu_a(l, \mathbf{s}, h) = f_a(l) + \gamma_a(\mathbf{s}) + \nu_a(h). \tag{3.6}$$

Here $f_a^l(l)$ is a continuous function of length, $\boldsymbol{\gamma}$ is a mean zero Gaussian spatial random field with Matérn covariance function, and $\boldsymbol{\nu}$ is an independent identically distributed Gaussian random haul effect. The spatial random field is intended to capture any spatial variation in the ALK. The haul random effect is intended to capture any haul variations, for example, a haul may by chance hit a school of fish of a certain age.

We assume that the spatially correlated Gaussian field in (3.6), $\boldsymbol{\gamma}$, follows a stationary Matérn covariance structure:

$$\mathrm{Cov}(\gamma(\mathbf{s}_1), \gamma(\mathbf{s}_2)) = \frac{\sigma_\gamma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa_\gamma||\mathbf{s}_1 - \mathbf{s}_2||)^\nu K_\nu(\kappa_\gamma||\mathbf{s}_1 - \mathbf{s}_2||), \tag{3.7}$$

where σ_γ^2 is the marginal variance, $||\cdot||$ is the Euclidean distance measure in kilometres, ν is a smoothing parameter, κ_γ is a spatial scale parameter and $K_\nu(\cdot)$ is the modified Bessel function of the second kind with $\nu = 1$. The spatial range parameter and marginal variances in the spatial fields are assumed to be equal across ages.

For each trawl haul, an ALK is obtained by maximizing the likelihood of the model in (3.5). The maximum likelihood estimate of μ_a is obtained using the R-package TMB (Kristensen et al., 2015) combined with the optimizing function *nlminb* in R. Advantages of using TMB in this application is that it utilizes the sparse structure in the precision matrix for the spatial field, it utilizes the Laplace approximation for the latent fields (both the spatial and the haul effect) for fast optimization of the hyperparameters, and it utilizes automatic derivation. Using such theory makes a good starting point for modeling of age distribution. A laptop with processor intel(R) Core(TM) i5-6300 CPU @ 2,40 GHz, used approximately 10 minutes to find the maximum likelihood estimate of the age given length model.

The spatial random field in the linear predictor for the age given length model (3.6) is estimated with the stochastic partial differential equation (SPDE) procedure described in (Lindgren et al., 2011). The theory behind the SPDE procedure is based on the precision matrix of a spatial field with Matérn covariance function can be approximated by a sparse matrix. This matrix is found by usage R-INLA package (Rue et al., 2009), and we further extracted the relevant parts needed from INLA to estimate the model in TMB.

3.3 Uncertainty estimation

We use nonparametric bootstrapping (Carpenter and Bithell, 2000) to estimate the uncertainty of age of estimated CPUEs. Four bootstrap procedures for simulating the data for uncertainty quantification are investigated: 1) a procedure suggested by DATRAS, which is based on hauls in the whole RFA, 2) a *stratified procedure*, which is similar to the DATRAS procedure but based on hauls in statistical rectangles, 3) *haul-based bootstrap procedure*, which accounts for the sampling variability in age-length compositions between hauls, 4) a *model-based ALK bootstrap procedure*, which allows a more objective and robust way of estimating ALK and accounts for the uncertainty that aries due to sampling variability.

3.3.1 DATRAS bootstrap procedure

The bootstrap procedure outlined by DATRAS (ICES 2006 or 2013) is as follows:

1. Assume there is n_{rec} trawl hauls in the ith statistical rectangle. Sample with replacement n_{rec} trawl hauls from the whole RFA and put them in the ith statistical rectangle.

2. Repeat step 1 for every statistical rectangle in the RFA.

3. Define $\mathbf{T}_{\text{sim}}^{\text{length}}$ to be the sample constructed with step 1-2.

4. Assume O_i is the number of age observations from ith length class in the RFA. Sample with replacement O_i of these observations. If there is only one observed age in that length class, sample either that fish or one which is closest in "length class distance".

5. Repeat step 4 for each length class with observed age.

6. Define $\mathbf{T}_{\text{sim}}^{\text{age}}$ to be the sample constructed with step 4-5.

7. Calculate the CPUE based on $\mathbf{T}_{\text{sim}}^{\text{length}}$ and $\mathbf{T}_{\text{sim}}^{\text{age}}$.

8. Repeat step 1-7 B times.

3.3.2 Stratified bootstrap procedure

We propose a stratified bootstrap approach, which is similar to the DATRAS bootstrap approach but which preserves both the number of trawl hauls within each statistical rectangle and the age observations within each length class. The stratified bootstrap procedure is as follows:

1. Assume there are $N_{\text{RFA}}^{(i)}$ trawl hauls in the ith statistical rectangle. Sample with replacement $N_{\text{RFA}}^{(i)}$ of the trawl hauls in the ith statistical rectangle. If there is only one trawl haul in the statistical rectangle, sample either that trawl haul or the closest in air distance.

2. Repeat step 1 for each statistical rectangle with trawl hauls.

3. Sample the catch-at-age (CA)-data with the same procedure as used in the DATRAS procedure.

4. Calculate CPUEs

5. Repetat step 1-4 B times.

Both the DATRAS and stratified bootstrap approaches sample age information in the whole RFA. However, given that the ALK is trawl dependent,that is, has a spatial structure on finer scale than the RFA, these procedures will underestimate the uncertainty.

3.3.3 Haul-based bootstrap procedure

1. Assume there are $N_{\text{RFA}}^{(i)}$ trawl hauls in the ith statistical rectangle. Sample with replacement $N_{\text{RFA}}^{(i)}$ of the trawl hauls in the ith statistical rectangle, and define $\mathbf{T}_{\text{sim}}^{\text{length}}$ to be that sample. If there is only one trawl haul in the statistical rectangle, sample either that trawl haul or the closest in air distance.

2. If there are no missing age-length compositions in the trawl hauls in the ith statistical rectangle calculate CPUEs.

3. If there are missing ages in the trawl hauls, then use the imputation procedure in Section C.2 in appendix C, and then calculate CPUEs.

4. Repeat steps 1-3 B times for the each statistical rectangle in the RFA

3.3.4 Model-based ALK bootstrap procedure

The bootstrap procedure used for calculating confidence intervals for the CPUE with use of the model-based ALK is constructed by sampling hauls stratified with respect to each statistical rectangle. The uncertainty in the ALK is taken into account by sampling from the joint normal approximation of the likelihood in each iteration in the bootstrap procedure. The joint precision matrix needed for the normal approximation is extracted from the estimated model in TMB.

4 RESULTS

Estimates of abundance indices are computed using 200 bootstrap replicates.

Table 3: Estimates of abundance indices for cod in RFA 7 in Q1 of year 2017. Estimated average standard error estimates (Se), and 95% confidence intervals (CI) for the simple, DATRAS and stratified bootstrap procedures are also given.

	Abundance indices		**Standard error for** $\mathrm{mCPUE}_{7,a}$			**Standard error for** $\mathrm{mCPUE}^*_{7,a}$		
Age (a)	$\mathrm{mCPUE}_{7,a}$	$\mathrm{mCPUE}^*_{7,a}$	Se_{sim}	Se_{DAT}	Se_{stra}	simple	DATRAS	stratified
1	0	0	0	0	0	0	0	0
2	2.316	2.316	0.759	0.909	0.454	0.759	0.909	0.454
3	4.262	4.262	1.532	2.218	0.836	1.532	2.218	0.836
4	2.023	2.023	0.685	0.732	0.440	0.685	0.732	0.440
5	1.769	1.769	0.816	0.819	0.626	0.816	0.819	0.626
6	1.124	1.124	0.491	0.477	0.316	0.491	0.477	0.316
7	0.355	0.355	0.271	0.176	0.157	0.271	0.176	0.157
				95% **CI from bootstrap procedures**				
1	0	0	0	0	0	(0,0)	(0, 0)	(0, 0)
2	2.316	2.316	(0.896, 3.612)	(0.692, 4.171)	(1.380, 3.231)	(0.896, 3.612)	(0.692, 4.171)	(1.380, 3.231)
3	4.262	4.262	(0.896, 3.612)	(0.692, 4.171)	(1.380, 3.231)	(0.687, 6.878)	(0.614, 8.975)	(2.651, 5.908)
4	2.023	2.023	(0.896, 3.612)	(0.692, 4.171)	(1.380, 3.231)	(0.945, 3.273)	(0.712, 3.437)	(1.176, 2.783)
5	1.769	1.769	(0.896, 3.612)	(0.692, 4.171)	(1.380, 3.231)	(0.544, 3.336)	(0.385, 3.270)	(0.640, 2.889)
6	1.124	1.124	(0.896, 3.612)	(0.692, 4.171)	(1.380, 3.231)	(0.462, 2.529)	(0.440, 2.259)	(0.660, 1.902)
7	0.355	0.355	(0.896, 3.612)	(0.692, 4.171)	(1.380, 3.231)	(0.066, 0.969)	(0.128, 0.737)	(0.138, 0.708)

5 DISCUSSION

Appendices

A Areas fished by different countries in the NS-IBTS

Typically, two different countries fish each rectangle so that at least two trawl hauls are made per rectangle. But, intensified sampling is carried out in the following areas: at least 3 hauls per rectangle are taken in statistical rectangles 31F1, 31F2, 32F1, 33F4, 34F2, 34F3, 34F4, 35F3, 35F4; while six or more hauls per rectangle are taken in statistical rectangles 30F1, 32F2, 32F3, 33F2, 33F3 (ICES 1999). The Skagerrak and Kattegat is fished solely by Sweden, who sample more than once in every rectangle while the west of Shetland (in Q1 and Q3) and inshore areas (Q3) is fished solely by Scotland. The edge of the Norwegian Trench is fished solely by Norway, but inshore areas near Denmark is fished by Denmark. The southern North Sea is fished by Denmark, Germany and England. France, typically, is the only country that surveys the western English Channel. Areas are surveyed by a single country because of the large proportion of untrawalable area (and subsequent gear damage issues experienced by other nations) for efficient logistical purposes.

B Otolith sampling per fish species

From 1991-2017, most countries conducted quota sampling of otoliths per length group in a RFA. But from 2013 Norway has been sampling one otolith per length class from each trawl haul (to 0.1cm below for shellfish, to 0.5cm below for herring and sprat and to 1cm below for all other species). From the first quarter in 2018 all countries are required to sample one otolith per length class per trawl haul. Table 4 gives the minimum sampling levels of otoliths for the target species. However, for the smallest size groups, that presumably contain only one age group, the number of otoliths per length class may be reduced, and more otoliths per length are required for the larger length classes.

Table 4: Minimum sampling levels of otoliths by species for RFA or per trawl haul.

Period	Species	Minimum sampling levels of otoliths per length class
1991-2017		**Number of otoliths per length class in a RFA**
	herring	8 otolihts per $\frac{1}{2}$ cm group
	sprat	16 otoliths per $\frac{1}{2}$ cm length class $8.0 - 11.0$ cm
		12 otoliths per $\frac{1}{2}$ cm length class ≥ 11.0 cm
	mackerel	8 otoliths per $\frac{1}{2}$ cm length class
	cod	8 otoliths per 1 cm length class
	haddock	8 otoliths per 1 cm length class
	whiting	8 otolihts per 1 cm length class
	Norway pout	8 otolihts per 1 cm length class
	saithe	8 otolihts per 1 cm length class
	All target species	from 2013 Norway has been sampling 1 otolith per length class from each trawl haul (to 0.1cm below for shellfish, to 0.5cm below for herring and sprat and to 1cm below for all other species).
2018		**Number of otoliths per length class per trawl haul**
	whiting	2 otoliths per 5 cm length class $11 - 15,\ 16 - 20,\ 21 - 25,\ 26 - 30$ cm
		2 otolihts per 1 cm length class > 30 cm
	Norway pout	2 otoliths per 5 cm length class $5 - 10,\ 11 - 15$ cm
		2 otolihts per 1 cm length class > 15 cm
	All other target species	1 otolith per length class (to 0.1cm below for shellfish, to 0.5cm below for herring and sprat and to 1cm below for all other species)

C Imputation for missing age samples

Catches of the target species are sampled (or subsampled with a size of 100 if the catches are too large) for length, and otoliths are typically collected from a subsample of the individuals sampled for length in the RFA, or per trawl haul as in the case of Norway for determining age of the fish (see Table 4). In the case of Norway where all trawl hauls are sampled for otoliths, missing age samples would still occur for the following two reasons: 1) the fish is below minimum length for otolith sampling (unreadable otoliths) or 2) otoliths are misplaced. Abundance indices by age group are estimated based on three age-length-keys (ALK): 1) DATRAS ALK estimator, 2) Haul dependent ALK estimator, and 3) Spatial model-based ALK estimator.

C.1 DATRAS ALK Borrowing Approach

The ALK proposed by DATRAS (ICES 2013), which is an aggregation of individual samples from a haul combined over a round fish area (RFA), and missing age samples are imputed as follows:

1. If there is no ALK for a length in the CPUE dataframe, age information is obtained accordingly
   - If length class (CPUE) $<$ minimum length class (ALK), then age=1 for the first quarter and age=0 for all other quarters
   - If minimum length class (ALK) $<$ length class (CPUE) $<$ maximum length (ALK) then age is set to the nearest ALK. If the ALK file contains values at equal distance, a mean is taken from both values.
2. If length class (CPUE) $>$ maximum length (ALK) age is set to the plus group.

The underlying assumption of this ALK approach is that age-length compositions are homogeneous within the superstrata.

C.2 Haul-based ALK Borrowing Approach

The second is an a haul dependent ALK estimator which we propose, and is denoted by ALK^H. Since the age-length composition of fish may be space-variant, that is, there may be variation in age-length compositions between trawl stations within a superstrata, the spatial dependence of the age-length composition must be accounted for to produce reliable estimates of the CPUE per age estimates. If this spatial dependence is ignored not only will estimates of abundance be biased but the impact on the variance may be substantial. So for each trawl haul an ALK^H is produced. Since there are few or none observations of ages for each length class in a trawl haul, length classes are therefore pooled in increasing order such that there are five length classes in each pooled length group. To replace missing values for the age distribution in the pooled length groups the method of ”borrowing” ages from length groups in trawl hauls closest in air distance within the RFA is used. If there are no observed ages in the pooled length group in the RFA, missing values for the age distribution are replaced following the procedure outlined in the DATRAS ALK procedure (C.1) in step 1.

References

Aanes, S. and Vølstad, J. H. (2015). Efficient statistical estimators and sampling strategies for estimating the age composition of fish. *Canadian journal of fisheries and aquatic sciences*, 72(6):938–953.

Agresti, A. (2003). *Categorical data analysis*, volume 482. John Wiley & Sons.

Berg, C. W. and Kristensen, K. (2012). Spatial age-length key modelling using continuation ratio logits. *Fisheries Research*, 129:119–126.

Carpenter, J. and Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? a practical guide for medical statisticians. *Statistics in medicine*, 19(9):1141–1164.

Dyke, G. and Patterson, H. (1952). Analysis of factorial arrangements when the data are proportions. *Biometrics*, 8(1):1–12.

Ehrhardt, N. M. and Legault, C. M. (1997). The role of uncertainty in fish stock assessment and management: a case study of the spanish mackerel, scomberomorus maculatus, in the us gulf of mexico. *Fisheries research*, 29(2):145–158.

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 lExploration De La Mer*, 86:1–5.

Fuller, W. A. (2011). *Sampling statistics*, volume 560. John Wiley & Sons.

Gerritsen, H. D., McGrath, D., and Lordan, C. (2006). A simple method for comparing age–length keys reveals significant regional differences within a single stock of haddock (melanogrammus aeglefinus). *ICES Journal of Marine Science*, 63(6):1096–1100.

Kimura, D. K. (1977). Statistical assessment of the age–length key. *Journal of the Fisheries Board of Canada*, 34(3):317–324.

Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. (2015). Tmb: automatic differentiation and laplace approximation. *arXiv preprint arXiv:1509.00660*.

Kvist, T., Gislason, H., and Thyregod, P. (2000). Using continuation-ratio logits to analyze the variation of the age composition of fish catches. *Journal of applied statistics*, 27(3):303–319.

Lehtonen, R. and Pahkinen, E. (2004). *Practical methods for design and analysis of complex surveys*. John Wiley & Sons.

Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498.

Ludwig, D. and Walters, C. J. (1981). Measurement errors and uncertainty in parameter estimates for stock and recruitment. *Canadian Journal of Fisheries and Aquatic Sciences*, 38(6):711–720.

Nottestad, L., Utne, K. R., 'Oskarsson, G. J., J'onsson, S. T., Jacobsen, J. A., Tangen, O., Anthonypillai, V., Aanes, S., Vølstad, J. H., Bernasconi, M., et al. (2015). Quantifying changes in abundance, biomass, and spatial distribution of northeast atlantic mackerel (scomber scombrus) in the nordic seas from 2007 to 2014. *ICES Journal of Marine Science*, 73(2):359–373.

Rindorf, A. and Lewy, P. (2001). Analyses of length and age distributions using continuation-ratio logits. *Canadian Journal of Fisheries and Aquatic Sciences*, 58(6):1141–1152.

Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392.

Walters, C. J. and Ludwig, D. (1981). Effects of measurement errors on the assessment of stock–recruitment relationships. *Canadian Journal of Fisheries and Aquatic Sciences*, 38(6):704–710.