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 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 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) which is assumed constant over relative large areas. In this paper we give a strong case for that the ALK varies within these areas by investigating a spatial model for the ALK. Because of the spatial variation we further investigate use of two different ALK which varies with space.

Figure 1 shows estimated age probabilities of a 30 cm long cod in the first quarter of 2014. The details of the model is given in section 3.1. As an example we can see that the age distribution for a 30 cm cod varied clearly in area mid and southern areas.

The objective of the paper is to compare the official reported CPUEs with similar CPUEs calculated with spatial varying ALK. We include two different spatial varying ALKs. One which is non parametric as

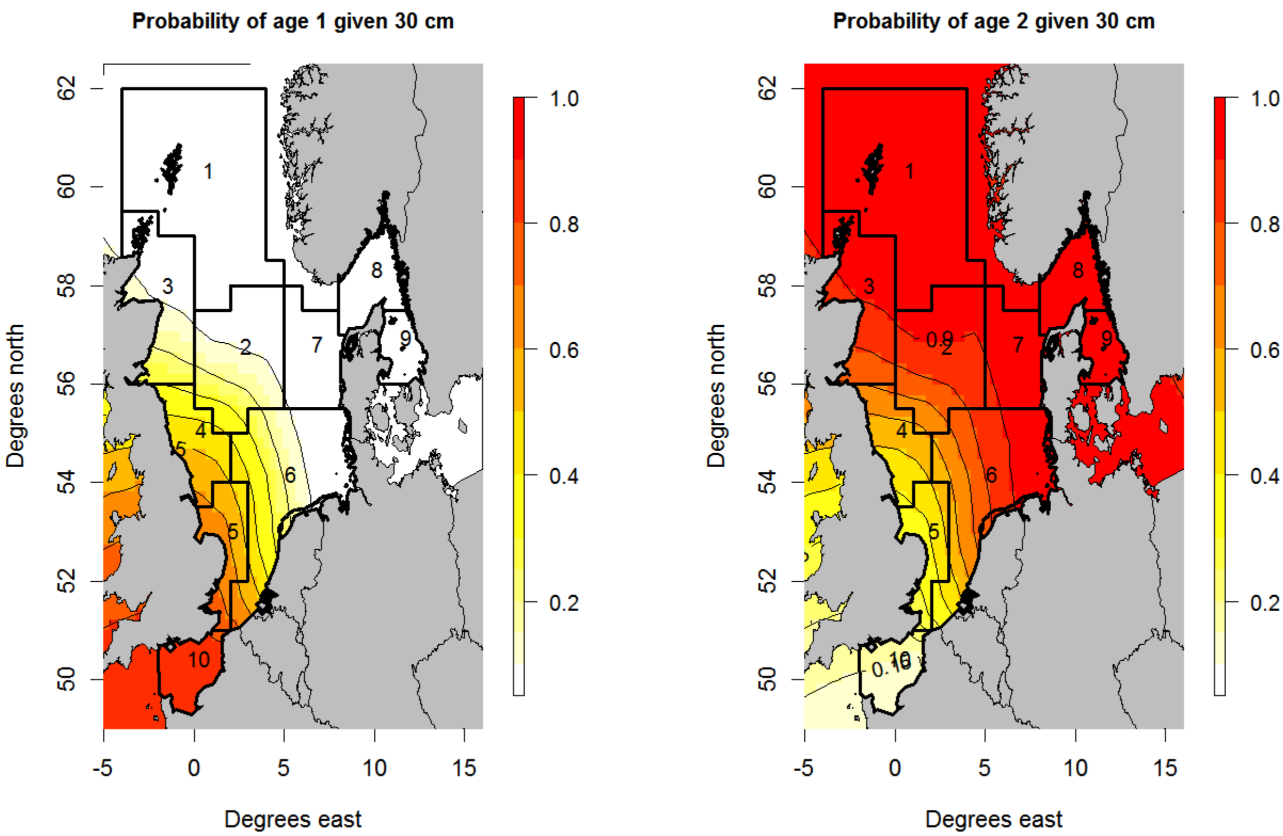

Figure 1: Estimated probability of age of a 30 cm long cod in the first quarter of year 2014. The probability of age three or older is approximately zero. The polygones marked 1 to 10 is the areas where the ALK is assumed constant in the currently used procedure for reporting official CPUEs.

used in the official estimates, and one spatial model based ALK.

The paper is organized as follows. Section 1.2 give an overview of the sampling procedure used by NS-IBTS. in section 3 we document the current procedure for calculating ALK and CPUE, and define our proposed ALKs. Section 5 documents the bootstrap procedures and the inference procedure for our model based ALK. In section 6 and 7 we include the results and discussion.

1.1 History 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 *Clu-*

pea 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 allow 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 scientific names (common 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 IBTS 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 (Table 6). The sampling frame is defined by the ICES index or roundfish areas (RFA) as shown in Figure 4, 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 3). 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 (Figure 2) 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: 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). **OLAV: SUGGEST TO MOVE ALL SUCH DETAILS TO THE APPENDIX, AND HAVE THIS SUBSECTION AS SHORT AND PRECISE AS WE MANAGE.** 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.

In principle, 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**. **Prior to 2012**, all countries, except England (in Q3) and Norway (in Q1 and Q3) randomly select hauling positions from a list of “clear” (and in many circumstances previously visited) haul positions. The same haul positions were used by Norway and England every year. However, **from 2012-2018** 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.

Table 2: Survey country, vessel name, and period research vessels participating in first quarter (Q1) and third quarter (Q3) during 1997-2017.

	First Quarter (Q1)		**Third Quarter (Q3)**	
Country	Vessel name	Period	Vessel name	Period
Denmark	Dana	January-February	Dana	July-August
France	Thalassa II	January-February	-	-
Germany	Walther Herwig III	January-February	Walther Herwig III	July-August
Netherlands	Tridens 2	January-February	-	-
Norway	G.O. Sars	January-February	Johan Hjort	July
UK England	-	-	Endeavour	August-September
UK Scotland	Scotia III	January-February	Scotia III	July-August
Sweden	Dana	January-February	Dana	August

1.2 Trawl Sampling and Protocols

The mulitpurpose chalut à Grande Ouverture Verticale (GOV) trawl (ICES 2012) is the recommended standard gear of the NS-IBTS and has been used on all participating vessels since 1992, while different pelagic and bottom trawls suitable for fishing finfish species were used before 1992. Since 1977, sampling of pelagic larvae during the International Bottom Trawl Survey in Q1 is also conducted using a standard Midwater Ring Net, commonly known as MIK. 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. GOV with standard groundrope with rubber discs (groundgear

A) for normal bottom conditions has been used throughout the survey area by all nations, except Scotland who since 1985 have used a hard ground gear for rough ground (groundgear B) on all stations north of 52° 30″ North (ICES 2012). During the tow it is imperative that the net geometry of the gear is within the acceptable limits for the depth of water (Figure 5). The trawls are towed in waters at a maximum depth of 200m in the North Sea and 250m in Division IIIa (Figure ....**insert figure showing map of NS-IBTS with RFA and divisions and hauls with age and lengths for a given year for example?**) with help of an "Exocet" kite and five floats attached to this kite. Rigging and trawl operation are described in (ICES 2012). The catching efficiency of the gear is assumed to be identical for every vessel. The tow duration was standardized to 30 minutes in 1978-2014 for all nations, except Scotland who maintained the tow duration of 60 minutes until 1998 (ICES 2015).

In the third quarter (Q3) of 2015, an experiment on tow duration of NS-IBTS hauls was conducted in the North Sea to investigate the effect on the composition of catches, and, which continued into the first quarter of 2016 (ICES 2015). **In this paper we have not consider the NS-IBTS dataset for these periods**.

Trawling is done during the day by all participating vessels from 2000-2017 while countries who did not participate in the sampling of herring larvae in Q1 trawled at night before 2000. Daylight hours are considered 15 minutes before sunrise to 15 minutes after sunset. 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 75 fish is selected. If a representative subsample cannot be selected further sorting of the species into two or more size grades or categories is necessary (ICES 2015). Otoliths are collected on board from a small fraction of all the target species from all RFA (Figure 4) to retrieve age reading. However, 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). Table 3 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 3: 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)

2 Data

Suggest we have a section with descriptive statistic of the IBTS data. E.g. The data used in this paper was collected in the first quarter of 2015. Table 4 gives an overview of the data used.

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). *(moved this from methods, Olav)*

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
Coordinates	Coordinates of each trawl haul
Location	Information of which statistical rectangle and RFA the haul belongs to
Duration of haul	Mean duration is 25.9 minutes, with 15 to 30 minutes as 90% coverage interval.

Table 4: *Summary of data NS-IBTS data for first quarter of 2015.*

3 METHODS

In this section we introduce the estimation procedrue for the ALK and CPUEs. First we define the CPUE estimate before we define the ALK used in the CPUE estimate.

3.1 CPUE estimates

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

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

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 defines as 3.2 and 3.3 with $\text{CPUE}_{h,l}$ substituted by $\text{CPUE}_{h,a}$. The purpose of this paper is to investigate different methods for calculating $ALK_{a,l,h}$. The different methods investigated is given in the next subsection.

ALK

In this subsection we document the different procedures for calculating the ALK investigated in this paper.

ALK used by DATRAS

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 borowing strength from neighbouring length classes is given in appendix A.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}.

Haul dependent ALK

We will now define a haul dependent ALK, denoted ALK^{H}, investigated in this paper. 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 typically are few observed length classes in a single haul **TODO, this is not implemented in our package yet**. 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 observed none 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 close in space is given in appendix A.2.

do we have overlapping of ages in grouped length bins in our ALK approach? If so bias would be introduced. According to Westrheim and Ricker (1978) ALK will have no bias only when ages do not overlap between length bins. OLAV: I AM NOT SURE WHAT YOU MEAN HERE, A FISH WITHIN A POOLED LENGTH CLASS CAN BE OF ANY AGE. THERE WILL TYPICALLY BE SEVERAL AGES WITHIN EACH POOLED LENGTH CLASS.

Spatial model based ALK

We will now define a model for the probabilities of a fish to be of age a given length, which is quite similar with the one introduced in (Berg and Kristensen, 2012). The ALK is then obtained by maximizing the likelihood of the model. Let $y(l, \mathbf{s}, h)$ be the age of a fish with length l, caught at location $\mathbf{s}$ by trawl haul h. Assume the probability of age a in a given year and quarter is given by:

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

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 Matern covariance function, and $\boldsymbol{\nu}$ is an iid Gaussian random haul effect. The spatial random field is intended to capture that there may be spatial variation in the ALK. The haul random effect is intended to capture that there may be haul to haul variations, e.g. a haul may by chance hit a shoal of fish of a certain age.

We assume that the spatially correlated Gaussian field in (3.6), $\boldsymbol{\gamma}$, follows a stationary Matern 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.

- **what is the estimator for age composition in the whole North Sea? - is it the average of** $\mathrm{mCPUE}_{p,a} = \sum_{l \in L} \mathrm{mCPUE}_{p,a,l}$ **in the 10 RFAs? and how is the variance computed?** *Olav: I also guess the average, and I have not seen any suggestion for how to calculate the variance of mCPUE in the whole North Sea. I would suggeste to combine all the bootstrap samples from all the RFAs.*

- **for the CPUE and CPUE* above if the ALKs are different, wouldn't the estimates be different? If both ALKs give the same estimates of the CPUE then we shouldn't distinguish between the two by calling one CPUE and the other** $CPUE^*$**, but instead just call the estimator CPUE?** *Olav: The estimated CPUE per age with different ALK will be different in all relevant cases.*

- **the stratified bootstrap procedure should also be different for the new ALK approach since it's at the haul level and not at the RFA? see step 4 in the stratified bootstrap procedure***Olav: I agree, I dont have any good suggestions right away. Perhaps sample from the statistical rectangle as done for the HL data. PERHAPS WE SHOULD JUST DROP THE SIMPLE AND STRATIFIED BOOTSTRAP PROCEDURE AND USE WHAT THEY SUGGEST (WHICH IS SOMETHING IN BETWEEN) AND COME UP WITH A BOOTSTRAP PROCEDURE SUITED FOR OUR ALK? FOR EXAMPLE COULD WE SAMPLE TRAWL HAULS STRATIFIED FOR EACH RECTANGLE AND THEN USE THEM BOTH FOR CONSTRUCTING THE ALK AND THE CPUE PER LENGTH.*

4 Uncertainty estimates

Olav: It gets a bit confusing with all the bootstrap procedures, perhaps it is best to discard the simple and the stratified, only use their suggestion and one bootstrap procedure suited for our new ALK procedure.

Uncertainty estimates of CPUEs is achieved by nonparametric bootstrap procedures (Carpenter and Bithell, 2000), and for the model based ALK the likelihood function is used directly for defining the uncertainty in the ALK. Three bootstrap procedures for simulating the data for uncertainty quantification are investigated: 1) A *simple procedure*, which is based on simple random sampling from the RFA. 2) A procedure suggested by DATRAS. 3) A *stratified procedure*, which is based on stratified/hierarchical sampling of the data.

4.1 Simple bootstrap

In this subsection we describe the simple bootstrap procedure used to quantify the uncertainty of the CPUE estimates in a given RFA. Assume there are N_{RFA} trawl hauls in the given RFA, where $N_{\text{RFA}}^{\text{age}}$ of them consists of age information. The simple bootstrap procedure is as follows:

1. sample with replacement N_{RFA} of the trawl hauls in the RFA, and define $\mathbf{T}_{\text{sim}}^{\text{length}}$ to be that sample.
2. Sample with replacement $N_{\text{RFA}}^{\text{age}}$ of the trawl hauls with age information and define $\mathbf{T}_{\text{sim}}^{\text{age}}$ to be the sample.
3. Calculate the CPUE based on $\mathbf{T}_{\text{sim}}^{\text{length}}$ and $\mathbf{T}_{\text{sim}}^{\text{age}}$. For uncertainty estimate using the model based ALK, the age is is sampled from the maximum likelihood estimate.
4. Repeat step 1-3 B times.

4.2 Stratified bootstrap

The IBTS try to sample trawl fish from every statistical rectangle and from every length class. Because of this we constructed the stratified bootstrap procedure in the following way:

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 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. 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 $i$te length class in the RFA. Sample with replacement O_i of these observations. If there are 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 $\mathbf{T}_{\text{sim}}^{\text{age}}$. For uncertainty estimate using the model based ALK, the age is is sampled from the maximum likelihood estimate.

8. Repeat step 1-7 B times.

4.2.1 Bootstrap similar as suggested by DATRAS

Here is what I interpret was datras suggest as a bootstrap procedure in the 2006 report. TODO:double check that this is what they mean.

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 recangle in the RFA.

3. Sample the trawl hauls with age data with the same procedure as used in the stratified procedure. It seems that datras suggest to merge length classes so that there is more then one observed fish inside each interval, but I don't find any clear documentation of what they think is the best way to merge length classes.

4. Calculate CPUEs

5. Repetat step 1-4 B times.

4.3 tmp

The estimators used for the NS-IBTS data are haul time-based for computing catch per unit effort (CPUE) indices. The indices are computed per roundfish area, which are specific for each species. Indices are computed as mean per stratum (statistical rectangle) and then as mean of the strata over the superstrata. In this paper we account for the uncertainty in abundance at age in the North Sea. Two estimators based on ALKs are considered to determine which estimator provides the most accurate estimates of precision given that the data are collected using a multistage sampling design. The first is an ALK, which is an aggregation of individual samples from a trawl haul combined over the round fish area (RFA) and which is the approach outlined by DATRAS. The second estimator uses an ALK method based on the trawl hauls, accounting for the variation in age-length composition between trawl hauls in a RFA. For this method, an ALK is produced for each trawl haul and abundance indices are estimated. The uncertainty in abundance at age is estimated using three bootstrap procedures: 1) a *simple nonparametric bootstrap* approach (Section 4.1), 2) *semi-stratified nonparametric bootstrap* proposed by DATRAS, but which has never been implemented. The second is a *stratified noparametric bootstrap* approach (Section 4.2), which accounts for the clustering effect in the multistage sampling design.

4.4 Imputation for missing age samples

OLAV: I SUGGEST TO MOVE ALL THIS TO THE APPENDIX, SEE A.1 and A.2.

Catches of the target species are sampled (or subsampled with a size of 75 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 3). 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 or 2) otoliths are misplaced. **OLAV: MISSING AGE INFORMATION OCCUR JUST BECAUSE THEY DO NOT TAKE**

OTHOLITS SAMPLE OF A FISH? I DID NOT UNDERSTAND THESE TWO REASONS. Abundance indices by age group are estimated based on two age-length-keys (ALK). The first is an 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. This is a rather strong assumption, and any violation would have serious 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. We therefore propose an ALK method that is based on trawl hauls, which we denote by ALK*. 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* 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 spatial 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 in step 1 above.

do we have overlapping of ages in grouped length bins in our ALK approach? If so bias would be introduced. According to Westrheim and Ricker (1978) ALK will have no bias only when ages do not overlap between length bins.

Table 5: List of symbols and parameters used.

Symbol	Definition
L	The set of length classes
A	The set of age groups
$\|A\|$	The number of age groups
P	The set of superstrata
$\|P\|$	The number of superstrata
S_p	The set of strata in superstrata p
$\|S_p\|$	The number of strata in p
H_s	The set of hauls in strata s
$\|H_s\|$	The number of hauls in s
U_h	The set of subsamples from haul h
$\|f_u\|$	The subfactor for the subsample u. The subfactor $\|f_u\|$ is always 1 for C-type data
$n_{u,l}$	The number of fish of target species in length class l in subsample u
d_h	The duration (minutes) for haul h
ALK	The age-length key for the target species **in a given population -further explanation?**.

5 Inference

The maximum likelihood estimates of the age given length model in section 3.1 is calculated with use of 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 SPDE procedure described in (Lindgren et al., 2011). The theory behind the SPDE procedure is based on that the precision matrix of a spatial field with Matern 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.

6 RESULTS

Estimates of abundance indices are computed using 200 bootstrap replicates.

Table 6: 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)

7 DISCUSSION

Appendices

A Missing ages

A.1 ALK^D

A.2 ALK^H

B

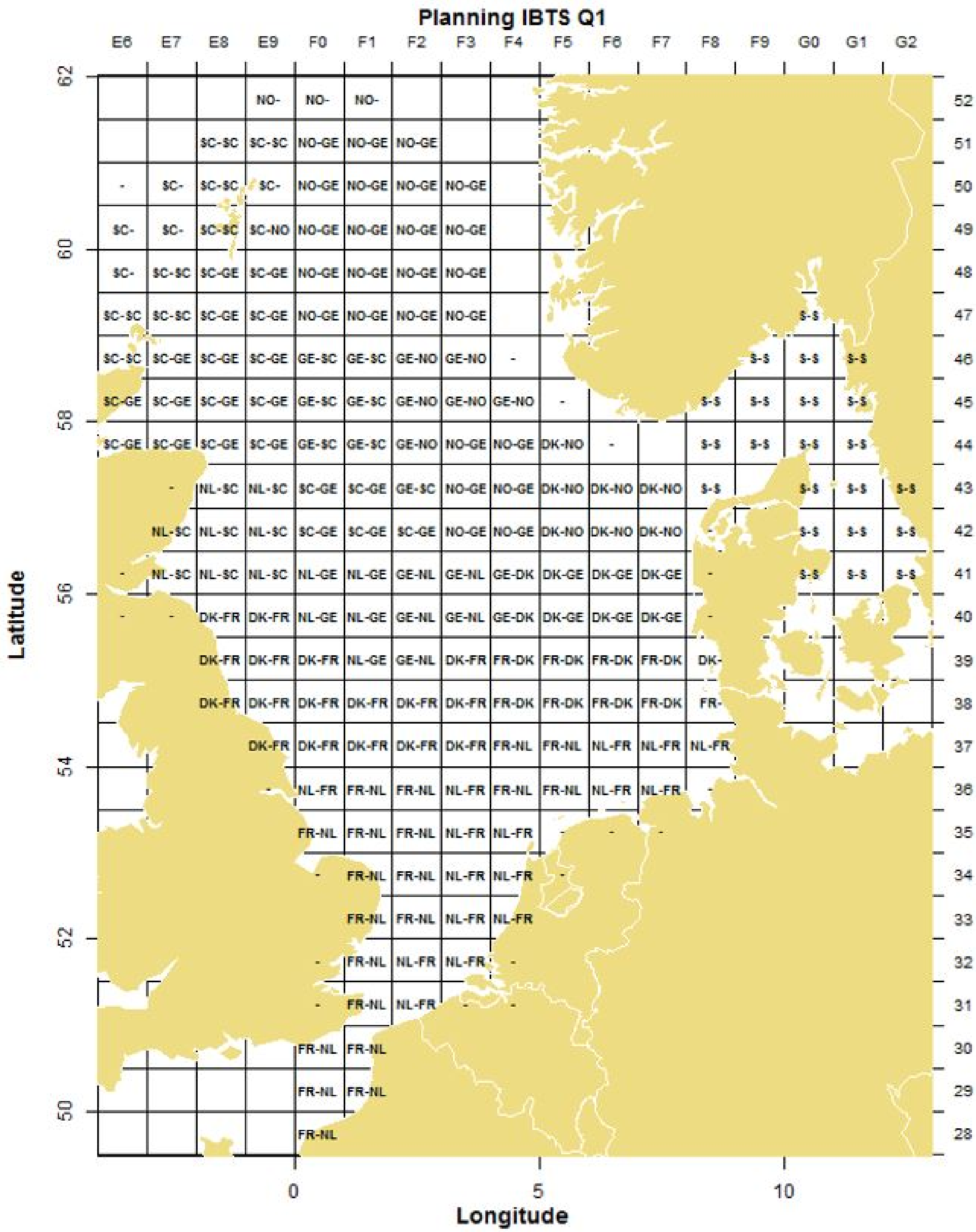

Figure 2: Spatial distribution of the ICES-rectangles in the IBTS Q1 over the participating countries. SC = Scotland, GE = Germany, NO = Norway, DK = Denmark, FR = France, NL = The Netherlands, S = Sweden (ICES 2016).

Weights of the statistical rectangle based on its surface area (10 – 200 meter in the North Sea and 10 -250 meter in the Skagerrak and Kattegat)

StatRec	Weight	StatRec	Weight	StatRec	Weight	StatRec	Weight	StatRec	Weight
31F1	0.6	38F0	1	41F6	1	44F1	1	47G0	0.3
31F2	0.8	38F1	1	41F7	1	44F2	1	47G1	0.02
31F3	0.05	38F2	1	41F8	0.1	44F3	1	48E6	1
32F1	0.8	38F3	1	41G0	0.2	44F4	1	48E7	1
32F2	1	38F4	1	41G1	0.97	44F5	0.9	48E8	0.9
32F3	0.8	38F5	1	41G2	0.53	44F8	0.25	48E9	1
32F4	0.01	38F6	1	42E7	0.4	44F9	0.8	48F0	1
33F1	0.3	38F7	1	42E8	1	44G0	0.94	48F1	1
33F2	1	38F8	0.3	42E9	1	44G1	0.6	48F2	1
33F3	1	39E8	0.5	42F0	1	45E6	0.4	48F3	0.5
33F4	0.4	39E9	1	42F1	1	45E7	1	48G0	0.02
34F1	0.4	39F0	1	42F2	1	45E8	1	49E6	0.8
34F2	1	39F1	1	42F3	1	45E9	1	49E7	1
34F3	1	39F2	1	42F4	1	45F0	1	49E8	0.4
34F4	0.6	39F3	1	42F5	1	45F1	1	49E9	1
35F0	0.8	39F4	1	42F6	1	45F2	1	49F0	1
35F1	1	39F5	1	42F7	1	45F3	1	49F1	1
35F2	1	39F6	1	42F8	0.2	45F4	0.6	49F2	1
35F3	1	39F7	1	42G0	0.32	45F8	0.3	49F3	0.5
35F4	0.9	39F8	0.4	42G1	0.89	45F9	0.02	50E6	0.1
35F5	0.1	40E7	0.04	42G2	0.64	45G0	0.24	50E7	0.6
36F0	0.9	40E8	0.8	43E7	0.03	45G1	0.55	50E8	0.7
36F1	1	40E9	1	43E8	0.9	46E6	0.4	50E9	0.9
36F2	1	40F0	1	43E9	1	46E7	0.9	50F0	1
36F3	1	40F1	1	43F0	1	46E8	1	50F1	1
36F4	1	40F2	1	43F1	1	46E9	1	50F2	1
36F5	1	40F3	1	43F2	1	46F0	1	50F3	0.2
36F6	0.9	40F4	1	43F3	1	46F1	1	51E6	0
36F7	0.4	40F5	1	43F4	1	46F2	1	51E7	0
36F8	0.5	40F6	1	43F5	1	46F3	0.8	51E8	0.5
37E9	0.2	40F7	1	43F6	1	46F9	0.3	51E9	1
37F0	1	40F8	0.1	43F7	1	46G0	0.52	51F0	1
37F1	1	41E6	0.03	43F8	0.94	46G1	0.2	51F1	1
37F2	1	41E7	0.8	43F9	0.41	47E6	0.8	51F2	0.5
37F3	1	41E8	1	43G0	0.21	47E7	0.6	51F3	0
37F4	1	41E9	1	43G1	0.7	47E8	1	52E6	0
37F5	1	41F0	1	43G2	0.3	47E9	1	52E7	0
37F6	1	41F1	1	44E6	0.5	47F0	1	52E8	0
37F7	1	41F2	1	44E7	0.5	47F1	1	52E9	0.1
37F8	0.8	41F3	1	44E8	0.9	47F2	1	52F0	0.2
38E8	0.2	41F4	1	44E9	1	47F3	0.6	52F1	0.5
38E9	0.9	41F5	1	44F0	1	47F9	0.01	52F2	0.1
								52F3	0

Figure 3

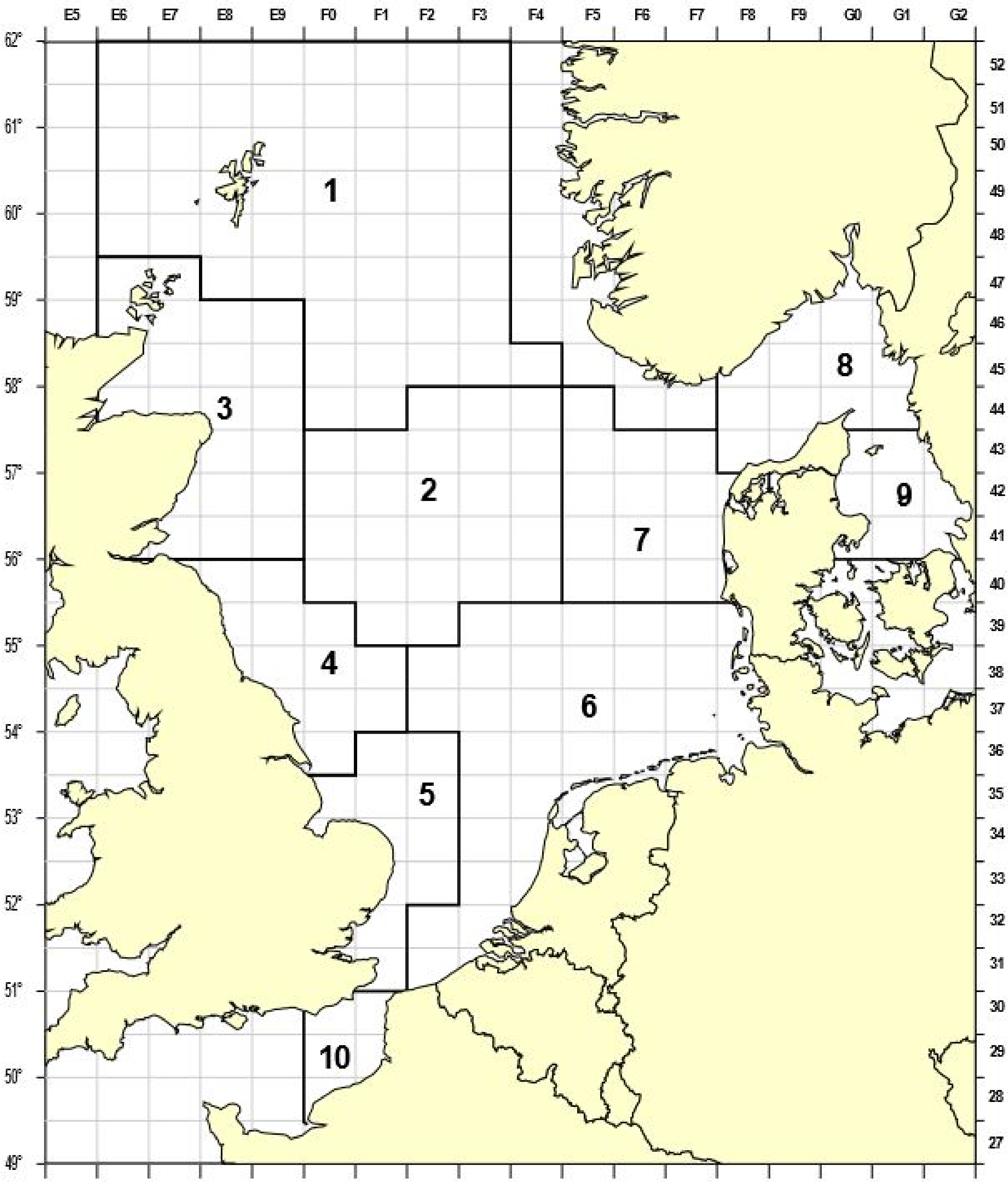

Figure 4: Standard roundfish areas used for roundfish since 1980, for all standard species since 1991. Additional RFA 10 added in 2009. For example, the number 1 indicates ICES Index Area 1, and an ICES Statitical rectangle (ST) in IA 1 is 43F1.

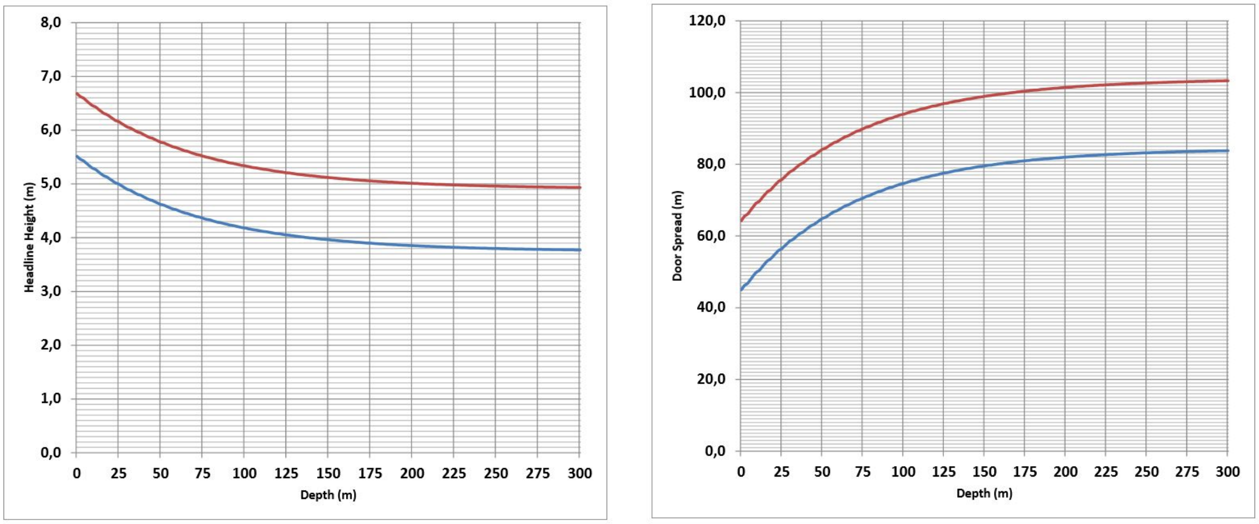

Figure 5: Left: Expected upper and lower limits of Headline height for water Right: Expected upper and lower limits of Door spread for water depth (ICES 2012).

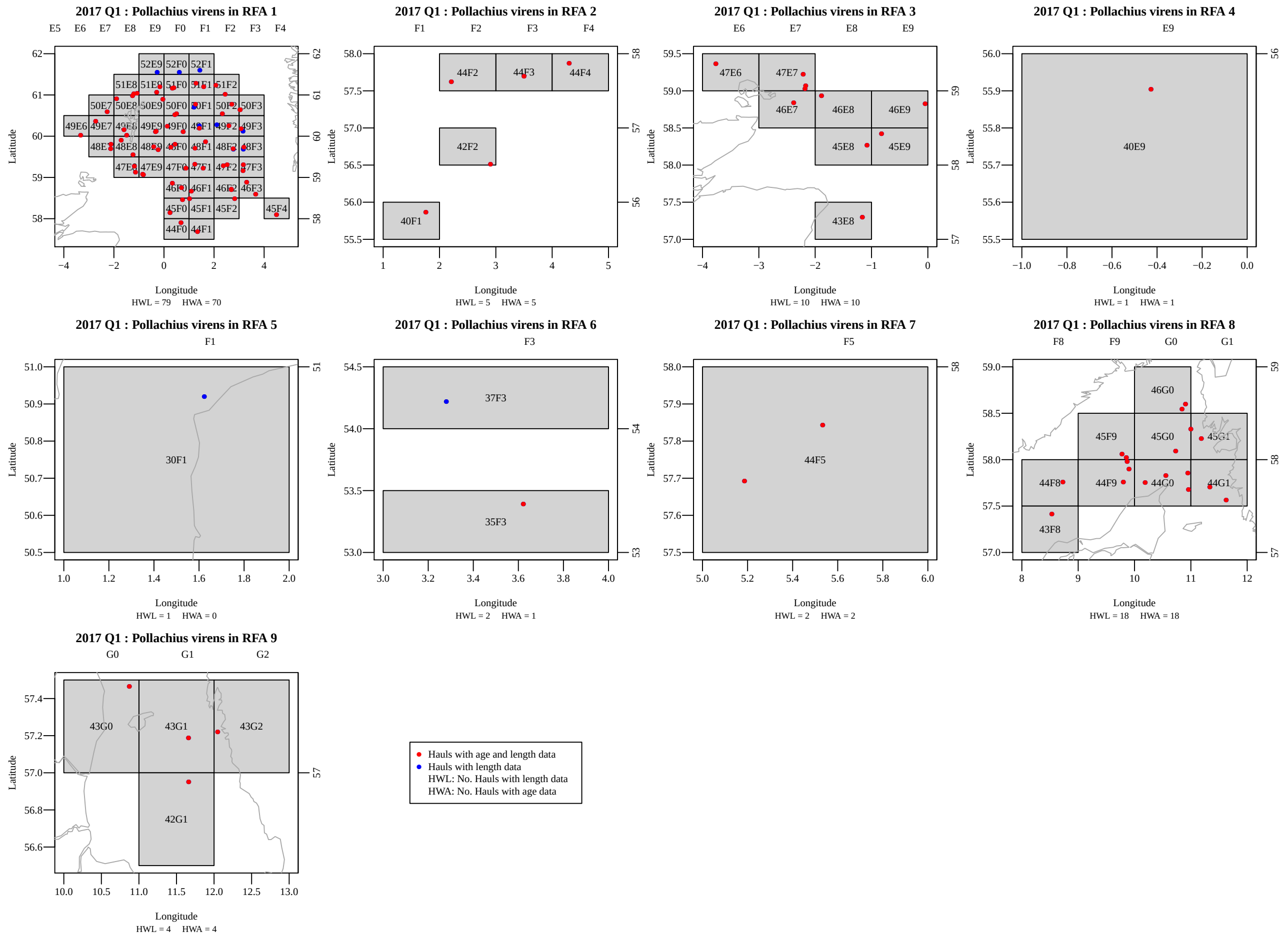

Figure 6: Plots of RFAs with trawl hauls having length and age information of Saithe in the first quarter of 2017.

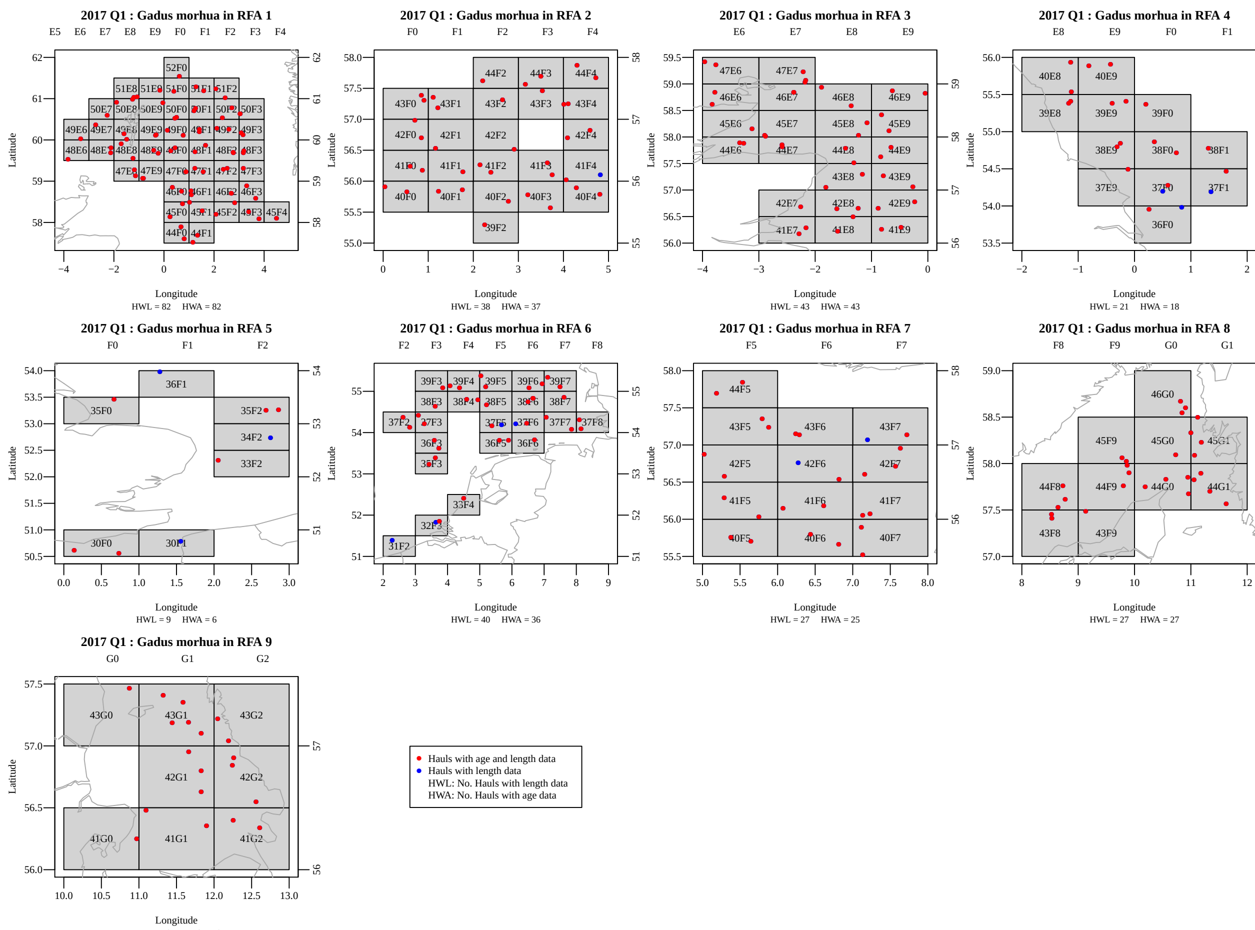

Figure 7: Plots of RFAs with trawl hauls having length and age information of Cod in the first quarter of 2017.

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.

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.

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.

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

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*.

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.

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.

Westrheim, S. and Ricker, W. (1978). Bias in using an age–length key to estimate age-frequency distributions. *Journal of the Fisheries Board of Canada*, 35(2):184–189.