

<sub>1</sub> An analysis of the North Sea International Bottom Trawl Survey

<sub>2</sub> Data

<sub>3</sub>

<sub>4</sub> **Abstract**

<sub>5</sub> In this research we present estimation procedures for calculating abundance at age indices, and inves-  
<sub>6</sub> tigate the sensitivity of these of the resulting estimates with respect to the number of otoliths collected at  
<sub>7</sub> sea. The procedures presented are applied to the North Sea International Bottom Trawls Survey data for  
<sub>8</sub> cod (*Gadus morhua*) and saithe (*Pollachius virens*). We demonstrate how much information would be  
<sub>9</sub> lost if the survey design was defined such that fewer otoliths were collected. Age length keys (ALKs) are  
<sub>10</sub> used to map lengths to age, and we use ALKs with and without the assumption of constant age length  
<sub>11</sub> structures over relatively large areas. All abundance at age indices are presented with variance estimates.

<sub>12</sub>

<sub>13</sub> **1 Introduction**

<sub>14</sub> Fish stock assessments are used by fishery managers for making management decisions regarding catch  
<sub>15</sub> quotas. The assessments provide fundamental information about the status of the stock, for instance,  
<sub>16</sub> whether the stock is increasing and support for increased levels of harvest should be given, or whether the  
<sub>17</sub> stock is decreasing and stricter control on harvest should be implemented. Associated with the parameters  
<sub>18</sub> used in fish stock assessment is their uncertainty, which should not be ignored when formulating management  
<sub>19</sub> policies (Walters and Ludwig, 1981; Ludwig and Walters, 1981; Berg et al., 2014). This uncertainty can arise  
<sub>20</sub> from many sources including natural variability, estimation procedures and lack of knowledge regarding the  
<sub>21</sub> parameter (Ehrhardt and Legault, 1997). The North Sea International Bottom Trawl Survey (IBTS) data,

22 coordinated by the International Council for the Exploration of the Sea (ICES), provides information on  
23 seasonal distribution of stocks and estimates of abundance indices and catch in numbers of fish per age-class  
24 without an assessment of the accuracy of these estimates. As stated by Ludwig and Walters (1981) it is  
25 relevant for managers to take into account the uncertainty related to stock size when making management  
26 polices. The indices of abundance at age from IBTS are based on data obtained from a stratified semi-random  
27 sampling approach of trawl stations, and it is essential to account for the sampling approach so as to produce  
28 reliable variance estimates (Lehtonen and Pahkinen, 2004). If the sampling approach is ignored, the effect on  
29 the variance of the parameters could be substantial. In particular, the variance could be greatly inflated due  
30 to the clustering effect, which involves intra-cluster correlation of the variables (Aanes and Vølstad, 2015;  
31 Lehtonen and Pahkinen, 2004).

32 There are two separate stages for generating abundance indices per age from the North Sea International  
33 Bottom Trawl Survey (IBTS) data. The first consist of calculating indices per *length* class, which are obtained  
34 by trawling in a stratified manner, sorting the catch by taxa and take biological measurement of the sorted  
35 catch. Then that knowledge is transformed to indices with respect to age. The latter part is achieved with  
36 an age-length key (ALK), which is constructed by sampling pairs of otoliths (each fish has a pair of otoliths)  
37 in a stratified procedure from each haul and/or sub-area. To our best knowledge, there has been no research  
38 on how much the uncertainty of the abundance indices is related to these two distinct parts. The main  
39 contribution of this research is to shed light on how the indices estimates and their associated uncertainty  
40 estimates change if less effort was spent on collection of otoliths. We achieve the reduction of otoliths by  
41 mimicking a defined sampling procedure with less effort. We also focus on the spatial distribution of the  
42 ALK, and such spatial structures in the ALK has also been investigated in Berg and Kristensen (2012) and  
43 Hirst et al. (2012).

44 Currently, abundance indices from IBTS are reported in DATRAS (ICES, 2018c) using an age-length key  
45 (ALK) (Fridriksson, 1934) which is assumed to be constant over relatively large areas. In this research we  
46 propose two ALKs which accounts for spatial variation: i) a nonparametric haul based ALK, and ii) a spatial  
47 model based ALK. These ALKs are described in Section 2. A spatial model based ALK (Berg and Kristensen,  
48 2012; Berg et al., 2014) known as the NS-IBTS Delta-GAM index (ICES, 2016b) is currently being used to

49 calculate standardized age-based survey indices used in assessment for the North Sea stock (haddock and  
50 cod). And, as far as we are aware the variance estimates of parameters estimated from NS-IBTS Delta-GAM  
51 index are *only* utilized for assessment of Herring (*Clupea harengus*) in the North Sea.

52 The spatial ALK model introduced in Berg and Kristensen (2012) is similar to the model used in this  
53 paper; the main difference is that we include the spatial structure through a spatial random field (Lindgren  
54 et al., 2011) and not through two dimensional splines (Wood, 2017).

55 An overview of the North Sea International Bottom Trawl Survey is given in Section 1.1. The current  
56 estimators for ALK and catch per unit effort (CPUE) used by ICES in their database for trawl surveys  
57 (DATRAS) and our proposed ALK estimators are given in Section 2. We apply these ALK methods to two  
58 case studies in Section 3, and a discussion is given in Section 4. R-code for reproducing the results can be  
59 fund at github (<https://github.com/natoyaj/TestPackage.git>).

## 60 **1.1 Overview of the North Sea International Bottom Trawl Survey**

61 The North Sea International Bottom Trawl Survey was formed in 1991, to combine the International Young  
62 Herring Survey (IYHS) and eight national surveys in the North Sea, Skagerrak and Kattegat areas. These  
63 surveys began in the 1960's, and the 1970's and 1980's, respectively. The IYHS was developed with the aim  
64 of obtaining annual recruitment indices for the combined North Sea herring (*Clupea harengus*) stock (ICES,  
65 2012), but yielded valuable information on other fish species such as cod (*Gadus morhua*) and haddock  
66 (*Melanogrammus aeglefinus*).

67 The North Sea IBTS began with quarterly surveys providing information on seasonal distribution of  
68 stocks sampled, hydrography and the environment, which allows changes in fish stock to be monitored and  
69 abundance of all fish species to be determined. These quarterly surveys, however became difficult to sustain  
70 as countries experienced budget cuts making it impossible to maintain high levels of research vessel effort. As  
71 such, in 1997 countries carried out a survey only twice a year; a first quarter survey (January-February) and  
72 a third quarter survey (July-September). The target species of IBTS fished from 1991-2018 includes standard  
73 pelagic species: Herring (*Clupea harengus*), Sprat (*Sprattus sprattus*) and Mackerel (*Scomber scombrus*); and  
74 standard roundfish species: Cod (*Gadus morhua*), Haddock (*Melanogrammus aeglefinus*), Saithe (*Pollachius*

*virens*), Norway Pout (*Trisopterus esmarkii*) and Whiting (*Merlangius merlangus*). There are also several by-catch species (see for example, ICES, 2006)

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, between 1997–2018 (Table S1.1 in Supplementary Materials S1 gives details of the research vessels). The sampling frame is defined by the ICES index or roundfish areas (RFA) as shown in Figure 1 numbered 1 to 10. These roundfish areas were substratified into small strata defined by non-overlapping statistical rectangles of roughly  $30 \times 30$  nautical miles ( $1^\circ$  Longitude  $\times$   $0.5^\circ$  Latitude), and were convenient to use for 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 (found in databases of national safe tows or DATRAS or commercial fishing data), 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 (ICES, 2018b). However, all countries select tows based on a semi-random approach from databases of national safe tows or DATRAS or commercial fishing data, except Sweden who uses fixed stations and in some cases depth-stratified semi-random sampling design (ICES, 2018b); and England who also uses fixed stations and only conduct surveys during the third quarter. 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 North Sea IBTS herring, saithe and sprat data are weighted by depth strata in the statistical rectangle (see Table S3.1 in Supplementary Materials S3). But this weighting is not included in the current estimation procedure in DATRAS. 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, 2015). But in some rectangles in the Eastern English Channel, Southern North Sea and Central North Sea intensified

102 sampling is carried out.

103 The recommended standard trawling gear of the North Sea IBTS is the mulitpurpose chalut à Grande  
104 Ouverture Verticale (GOV) trawl (ICES, 2012), which has been used on all participating vessels since 1992,  
105 while different pelagic and bottom trawls suitable for fishing finfish species were used before 1992. Standard-  
106 ized trawling protocols were adopted with a towing speed of 4 knots but depending on vessel performance,  
107 tide and weather conditions the average towing speed can be at minimum 3.5 and maximum 4.5 knots. From  
108 2000-2018 trawling is done during the daylight hours, which are considered 15 minutes before sunrise to 15  
109 minutes after sunset (ICES, 2012). After each trawl the total catch of the different species is weighed on  
110 board and biological parameters such as length for all fish species caught (to 0.1 cm below for shellfish, to  
111 0.5 cm below for herring and sprat and to 1 cm below for all other species) are collected. Where the numbers  
112 of individuals are too large for all of them to be measured to obtain the length distribution, a representative  
113 subsample of 100 fish is selected. A pair of otoliths are collected on board from a small fraction of all the  
114 target species from all RFAs (Figure 1) to retrieve age reading. Table S2.1 in Supplementary Materials S2  
115 gives the minimum sampling levels of otoliths for the target species.

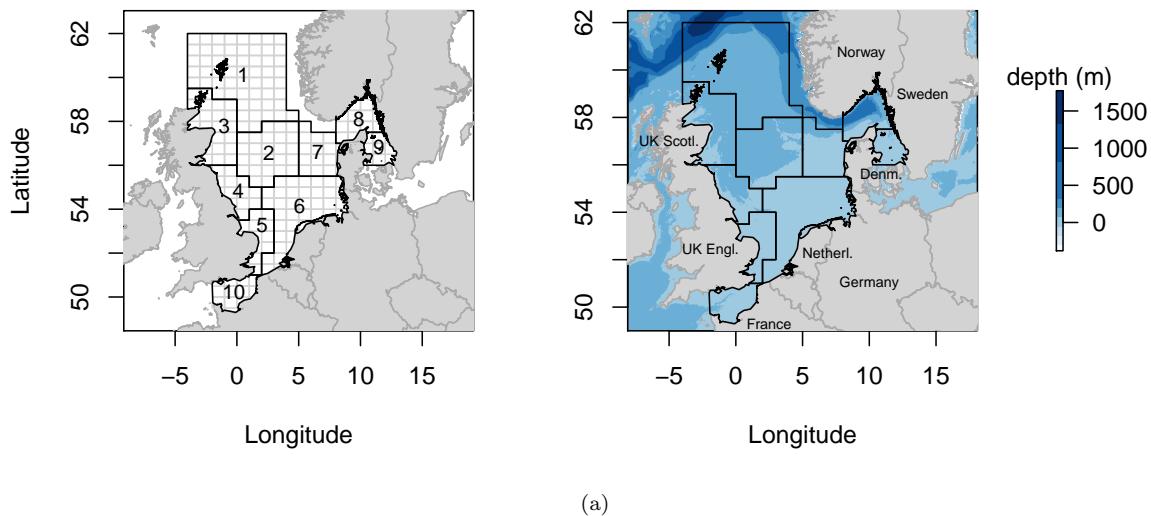


Figure 1: Standard roundfish areas (RFAs) used for roundfish since 1980 and for all standard species since 1991 (left panel). RFA 10 was added in 2009. The number 1, for example, indicates ICES RFA 1. The small grey rectangles in the left panel indicates the statistical rectangles of approximately 30 × 30 nautical miles (these vary from 28 nm wide in the north, to 40 nm wide in the south of North sea) ( $1^{\circ}$  Longitude ×  $0.5^{\circ}$  Latitude). The map in the right panel shows the Norwegian trench and shelf edge (depths 1000-1500).

## 2 METHODS

117 This section gives the estimators of abundance indices. The estimators are haul-duration based and utilizes  
 118 an ALK approach. We consider the ALK approach used in DATRAS and we propose two ALK estimators.  
 119 The ALK used in DATRAS for computing abundance indices does not account explicitly for the spatial  
 120 distribution in age-length structures over large areas. As differences in age-length structures may exist over  
 121 large areas, these differences do have the potential to result in a biased ALK (Gerritsen et al., 2006; Kimura,  
 122 1977). These differences may be caused either by variation in length-at-age distributions or by variations in  
 123 the relative abundance of age classes, that is age-at-length distributions (Gerritsen et al., 2006). To account  
 124 for the spatial distribution we propose a design-based ALK estimator that is haul dependent (Section 2.2.2)  
 125 and a model based ALK estimator (2.2.3).

126 ***2.1 Catch per unit effort***

127 In this research, the catch per unit effort (CPUE) is defined as the number of fish of a certain species and  
 128 age or length which are caught per hour trawl. In this section we define the CPUE mathematically, which  
 129 explains how the index is calculated. For a given species of interest, let  $n_{h,l}$  be the number of fish with  
 130 length  $l$  caught by trawl haul  $h$ . The CPUE for a given length  $l$  by trawl haul  $h$  is defined as

$$\text{CPUE}_{h,l} = \frac{n_{h,l}}{d_h}, \quad (2.1)$$

131 where  $d_h$  is the duration of the trawl in hours. The CPUE per age class is further defined as

$$\text{CPUE}_{h,a} = \sum_{l \in \mathbf{L}} \text{CPUE}_{h,l} \times \text{ALK}_{a,l,h}, \quad (2.2)$$

132 where  $\mathbf{L}$  is the set of all length classes and  $\text{ALK}_{a,l,h}$  is the age length key, which represents the estimated  
 133 proportion of fish with age  $a$  in  $l$ th length class in haul  $h$ . For a given number of trawl hauls in a statistical  
 134 rectangle, the mean CPUE defined as mCPUE in a statistical rectangle can be expressed as the average  
 135 CPUE of the trawl hauls in the statistical rectangle:

$$\text{mCPUE}_{s,a} = \frac{\sum_{h \in H_s} \text{CPUE}_{h,a}}{|H_s|}. \quad (2.3)$$

<sup>136</sup> Here  $H_s$  represents the set of trawl hauls taken in statistical rectangle  $s$ , and  $|H_s|$  is the number of hauls  
<sup>137</sup> taken in the rectangle. The mCPUE in  $p$ th RFA is further defined as

$$\text{mCPUE}_{p,a} = \frac{\sum_{s \in S_p} \text{mCPUE}_{s,a}}{|S_p|} \omega_s, \quad (2.4)$$

<sup>138</sup> where  $S_p$  is the set of all statistical rectangles in RFA  $p$ ,  $|S_p|$  is the number of statistical rectangles in RFA  $p$ ,  
<sup>139</sup> and  $\omega_s$  is a weight factor for each statistical rectangle (see Table S3.1 in Supplementary Materials S3). For  
<sup>140</sup> species such as saithe, herring, and sprat the indices at age are calculated using the mean over rectangles,  
<sup>141</sup> weighted for the percentage of area with water depths between 10m-200m, and for RFAs 8 and 9 water  
<sup>142</sup> depths between 10m-250m (ICES, 2013a).

<sup>143</sup> The mean catch per unit at age in the whole study area is defined as

$$\lambda_a = \frac{\sum_{p \in \mathbf{P}} A_p \text{mCPUE}_{p,a}}{A_{\text{total}}}. \quad (2.5)$$

<sup>144</sup> We refer to (2.5) as the index of abundance at age, where  $\mathbf{P}$  is the set of RFAs,  $A_p$  is the area of RFA  $p$ ,  
<sup>145</sup> and  $A_{\text{total}} = \sum_{p \in \mathbf{P}} A_p$ .

## <sup>146</sup> 2.2 **ALK estimators**

<sup>147</sup> The definition of the CPUE of age includes an ALK, see (2.2), which we described in this section. Three  
<sup>148</sup> ALK estimators are included in this research, which are named as follows: *i*) DATRAS ALK, *ii*) haul based  
<sup>149</sup> ALK and *iii*) model based ALK.

### <sup>150</sup> 2.2.1 *Area based ALK*

<sup>151</sup> We refer to the ALK used in DATRAS to estimate abundance at age for the IBTS data as an area based  
<sup>152</sup> ALK, defined as  $ALK_{a,l,h}^A$ . The area based ALK is defined as constant within each RFA, and is calculated  
<sup>153</sup> for each RFA by aggregating the age observation from each RFA.  $ALK_{a,l,h}^A$  used in equation (2.2) is defined  
<sup>154</sup> as the proportion of observed fish with age  $a$  in length class  $l$  in the RFA  $h$ . If there are no observed fish  
<sup>155</sup> in length class  $l$  in the RFA, ages from length classes close to  $l$  is used. The details of the procedure for  
<sup>156</sup> borrowing age data from neighbouring length classes are given in Supplementary Materials S4.1. We want  
<sup>157</sup> to highlight that missing ages occur seldom, in e.g. quarter 1 year 2018, 99.1% of the cod had an observed

158 age in the corresponding RFA within the same cm group. The underlying assumption of this ALK is that  
 159 age-length compositions are homogeneous within the RFAs. This is a rather strong assumption, and any  
 160 violation would have an unknown impact on the estimates of abundance indices. Aanes and Vølstad (2015)  
 161 illustrated that violation of the assumption of constant ALK leads to biased estimates of CPUEs.

162 *2.2.2 Haul based ALK*

163 We define a haul dependent ALK by  $ALK^H$ . The  $ALK_{a,l,h}^H$  used in equation (2.2) is defined as the average  
 164 proportion of observed fish with age  $a$  in length class  $l$  in haul  $h$ . If there are no observed ages of fish in a  
 165 length class  $l$  in the haul, ages from the same length class in the haul close by is used (see Supplementary  
 166 Materials S4.2 for the procedure). We want to highlight that missing ages occur seldom, in e.g. quarter 1  
 167 year 2018, 98.6% of the cod had an observed age in the corresponding haul within the same cm group.

168 *2.2.3 Model based ALK*

169 In this section we introduce a spatial model based ALK, which we define as  $ALK^M$ . Using such a model  
 170 enables us to obtain smooth structures in the distribution of age given length. It further enables us to utilize  
 171 spatial latent effects. Spatial model based approach of age-lengths are widely used (Berg and Kristensen,  
 172 2012; Hirst et al., 2012; Rindorf and Lewy, 2001), and are used for stock assessment in the North Sea (Berg  
 173 et al., 2014).

174 Let the response variable of the age group of a fish be  $a = M, \dots, A$  where  $M$  is the youngest age, and  $A$   
 175 is the oldest age which is typically defined as a "plus group". Suppose  $y(l, \mathbf{s})$  is the age of a fish with length  
 176  $l$  caught at location  $\mathbf{s}$ . As in Berg and Kristensen (2012) we use a continuous ratio model for the spatial age  
 177 given length model. However, in our application we assume for each species we know a length  $l^*$  such that  
 178 all fish above length  $l^*$  are above age  $M$ , and all fish with length below  $l^*$  are of age below  $A$ . By including  
 179 such a variable we reduce the number of parameters in the model by removing one linear predictor. Define  
 180 the continuous ratio we are modelling as

$$\pi_a[y(l, \mathbf{s})] = \frac{p_a(l, \mathbf{s})}{p_a(l, \mathbf{s}) + \dots + p_A(l, \mathbf{s}) + p_M(l, \mathbf{s})} \quad \text{for } a = M + 1, \dots, A - 1, \quad (2.6)$$

181 where  $p_a(l, \mathbf{s})$  is the probability of a fish with length  $l$  at location  $\mathbf{s}$  to be of age  $a$ . Note that either  $p_A(l, \mathbf{s})$

182 or  $p_M(l, \mathbf{s})$  is known to be equal to zero, and the other is selected such that  $\sum_a p_a = 1$ . We assume the logit

183 link

$$\log \left[ \frac{\pi_a[y(l, \mathbf{s})]}{1 - \pi_a[y(l, \mathbf{s})]} \right] = f_a(l) + \gamma_a(\mathbf{s}), \quad (2.7)$$

184 where  $f_a(l)$  is a continuous function of length and  $\gamma$  is a mean zero Gaussian spatial random field with  
185 Matérn covariance function (Stein, 2012). The spline  $f_a(l)$  is intended to account for the fact that longer  
186 fish are typically older, and the spatial random field,  $\gamma$ , is intended to account for spatial variation in the  
187 ALK. The continuous function  $f_a(l)$  in (2.7) is modelled with usage of P-splines (Wood, 2017), and these  
188 spline regression coefficients are included as a mean zero Gaussian random effect. The precision matrix for  
189 the spline regression coefficients is constructed such that wiggleness is penalized, see Wood (2017, page 239)  
190 for details. The R package mgcv (Wood, 2015) is used for extracting the precision matrix needed for the  
191 spline regression coefficients. The marginal variance of the P-splines regression coefficients,  $\sigma_f^2$ , is estimated  
192 in our inference procedure.

193 We assume that the spatially Gaussian random field in (2.7),  $\gamma$ , follows a stationary Matérn covariance  
194 structure defined as

$$\text{Cov}(\gamma(\mathbf{s}_1), \gamma(\mathbf{s}_2)) = \frac{\sigma_\gamma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa\|\mathbf{s}_1 - \mathbf{s}_2\|)^\nu K_\nu(\kappa\|\mathbf{s}_1 - \mathbf{s}_2\|), \quad (2.8)$$

195 where  $\sigma_\gamma^2$  is the marginal variance of the spatial field;  $\|\mathbf{s}_1 - \mathbf{s}_2\|$  is the distance between  $\mathbf{s}_1$  and  $\mathbf{s}_2$  in kilometres;  
196  $\kappa$  is a spatial scale parameter of the spatial field;  $\nu$  is a smoothing parameter and  $K_\nu(\cdot)$  is the modified Bessel  
197 function of the second kind with  $\nu = 1$ . The spatial field is estimated with the stochastic partial differential  
198 equation (SPDE) procedure described in Lindgren et al. (2011). The main concept behind the SPDE  
199 procedure is that the precision matrix of a spatial field with Matérn covariance function can be approximated  
200 by a sparse matrix on a grid covering the area of interest. Such a grid and sparse precision matrix are  
201 constructed with use of the R-INLA package (Rue et al., 2009). Figure 2 gives an illustration of the mesh used  
202 for approximating (2.8) for one of our focal species (cod), which we describe in Section 3. Detail regarding the  
203 construction of the mesh can be found at github (<https://github.com/OlavNikolaiBreivik/IBTSspatialALK>).

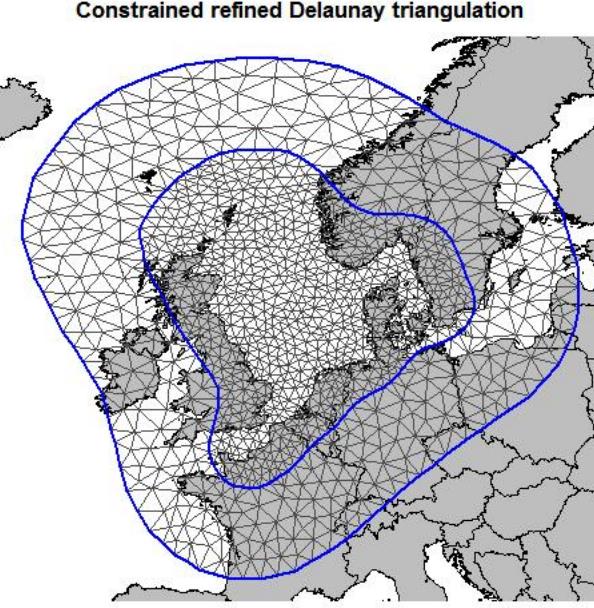


Figure 2: Mesh used in the case study for cod in Section 3 for approximating (2.8) with the SPDE-procedure. The species specific constant  $l^*$  is selected as the mid point between the shortest fish of age A and the longest fish of age M in the corresponding year and quarter. A sensitivity analysis of this constant were performed by adjusting it up and down 5 cm for cod in year 2018 in Q1. The point estimate of the mCPUEs then changed in the forth decimal, which we will consider negligible.

204     The model based ALK estimate is obtained by maximizing the likelihood. We maximize the likelihood  
 205    with use of an R-Package called Template Model Building **TMB** (Kristensen et al., 2015), combined with the  
 206    optimizing function **nlnmb** in R. In this application **TMB** is advantageous as it uses Laplace approximation  
 207    for the latent fields gaining computational efficiency, it also utilizes sparse structures in the latent fields, and  
 208    uses automatic derivation.

### 209    2.3    *Uncertainty estimation*

210    In this section we describe how the uncertainty of the CPUE estimates are calculated. We use nonparametric  
 211    bootstrapping to quantify the uncertainty of the CPUEs. In nonparametric bootstrapping independent sam-  
 212    ples of lengths and age are drawn with replacement from the original data and approximate 95% confidence  
 213    intervals are obtained using bias-corrected percentile method (Carpenter and Bithell, 2000). Nonparametric  
 214    resampling allows us to estimate the sampling distribution of the CPUE empirically without making as-  
 215    sumptions concerning the data. The bias-Corrected method adjusts for the bias and skew of the sampling  
 216    distribution of the data (Puth et al., 2015; Karlsson, 2009). The bootstrap procedure is given in Supple-

217   mentary Materials S5. (The procedure in S5 is a general description of a bootstrap procedure, this should  
218   be just included with a reference.)

219   A bootstrap procedure for estimating the uncertainty of CPUEs in the North Sea is suggested in ICES  
220   (2006). This procedure is given in Supplementary Materials S5 (The procedure in S5 is a general description  
221   of a bootstrap procedure, this should be just included with a reference.). In the rest of this research, we  
222   refer to this procedure as DATRAS bootstrap procedure, as it is the current procedure outlined in DATRAS  
223   for uncertainty estimation of IBTS indices. However, this bootstrap procedure has never been implemented  
224   in DATRAS. The DATRAS procedure is divided into two parts; one part which samples CPUE per length  
225   (2.1), and another part which samples the ALK used in (2.2). The DATRAS bootstrap procedure is based  
226   on the assumption of homogeneous CPUE within RFAs. This assumption is likely to be wrong, and would  
227   typically cause an overestimation of the uncertainty. Therefore, we have included a bootstrap procedure,  
228   defined as the stratified bootstrap procedure, which instead assumes constant CPUE within each statistical  
229   rectangle.

230   2.3.1   *Bootstrap procedures and reducing effort*

231   In this section we describe the two bootstrap procedures used in this research. The two procedures differ by  
232   how the ages are sampled. The DATRAS and stratified bootstrap procedure are constructed as follows:

233   1. For each statistical rectangle, sample  $|H_s|$  hauls with replacement from the corresponding statistical  
234   rectangle. If there is only one haul within a statistical rectangle, sample either that haul or the closest  
235   haul.

236   2. Sample age observations from the simulated hauls obtained by step 1. For the DATRAS procedure  
237   this is achieved by (a), and for the stratified procedure this is achieved by (b):

238   (a) For each RFA and length group, sample with replacement  $n_{l,a,RFA}$  age observations stratified  
239   with respect to RFA and length group. Here  $n_{l,a,RFA}$  is the total number of age observations in  
240   length group  $l$  in the corresponding RFA. If there is only one observed age from a given length  
241   group, i.e.  $n_{l,a,RFA} = 1$ , we sample either that age or an age in the closest length class with  
242   observed ages within the RFA.

243 (b) For each *haul* and length group, sample without replacement  $n_{l,a,h}$  age observations stratified  
244 with respect to *haul* and length group using a pseudo-population bootstrap procedure (Mashreghi  
245 et al., 2016) (TODO). Here  $n_{l,a,h}$  is the total number of age observations in length group  $l$  in the  
246 corresponding haul. Ages are sampled with probability proportional to the number of fish in the  
247 corresponding cm group and haul. If there is only one observed age within a length group in a  
248 haul, that age is sampled.

249 3. Calculate mCPUE as explained from equation (2.1) to (2.5) with use of the sampled data.

250 4. Repeat 1-3 B times.

251 In this research we investigate how the estimated mCPUE is affected by reducing the number of age  
252 readings. We simulated reduction of otoliths by choosing  $n_{l,a,h}$  in step 2b equal to a pre defined number.  
253 In our case study we have chosen  $n_{l,a,h}$  equal *one* when simulating the reduced effort. By doing that, the  
254 simulated data sets in the bootstrap procedure are possible realisations obtained by collecting only *one*  
255 otoliths per length group at sea.

### 256 3 Case studies

257 In this section we apply the methods described in Section 2 to data for cod and saithe from the International  
258 Bottom Trawl Survey for the years 2014-2018, which is obtained from the DATRAS database (ICES, 2018c).  
259 These years are chosen for several reasons. The first is that in year 2018 new sampling procedures proposed  
260 by ICES for the collection of otoliths were introduced in the surveys. For instance, *one* pair of otoliths  
261 per length group is sampled for most target species (see Table S2.1 in Supplementary Materials S2 for  
262 the sampling procedures for each target species), and this data is appropriate for the application of our  
263 proposed sample optimization procedure described in Section ???. The second is that IBTS included age 0  
264 in Q3 surveys. Also, some species such as saithe that occupies the deeper waters in the northern part of  
265 the North Sea and in the Skagerrak and Kattegat, along the shelf edge (ICES, 2018a), the IBTS Q3 data  
266 is relevant for analyses compared with data from IBTS Q1 surveys, which do not adequately cover these  
267 areas where saithe is distributed (ICES, 2016a); see Figure 1. In the third quarter of 2015, an experiment

268 on tow duration of the North Sea IBTS hauls was conducted to investigate the effect on the composition  
269 of catches, and which continued into the first quarter of 2016 (?). A mix of 15 minutes and 30-minutes  
270 hauls were used in all rectangles to which two hauls have been allocated: one 15-minutes haul and one 30-  
271 minutes haul maintaining a full North Sea-wide set of regular 30-minutes hauls in case there were significant  
272 differences in the 15-minutes haul. Four nations (Denmark, Germany, Norway and Scotland) participated  
273 in the experiment, while two nations (England and Sweden) retained haul duration for all of their hauls at  
274 30-minutes. France conducted some additional 15-minutes hauls, paired with 30-minutes hauls during the  
275 first quarter of 2016 (?). These years provide extensive sample data for cod and saithe with yearly sampling  
276 effort varying between 333 and 389 trawl hauls. For each of these trawl hauls sampled in the years 2014-2018,  
277 concurrent length measurements and otoliths for age determination were obtained from a subsample of cod  
278 and saithe. The subsample of otoliths taken for age determination varied between 1600 and 2895 for cod,  
279 and between 581 and 1631 for saithe. Table S6.1 in Supplementary Materials S6 briefly describes the data  
280 for the years 2014-2018.

### 281 ***3.1 Estimated indices of abundance and variability for cod and saithe***

282 In this section we apply the three ALK methods given in section 2.2 for abundance estimation, and the  
283 bias-corrected bootstrap method, given in Section 2.3.1 for estimating variability of estimated indices of  
284 abundance. In this section we apply the three ALK methods, given in section 2.2, for estimating abundance  
285 at age and the bias-corrected bootstrap method, given in Section 2.3.1, for estimating variability of estimated  
286 indices of abundance. The main assumption of the area based ALK is that the age-length compositions of  
287 species over large areas are the same. Figure 3 illustrates the predicted probability of age of cod given length  
288 using the spatial model based ALK (2.7). Figure 3 illustrates that the main assumption of the area based  
289 ALK of constant age-length compositions over large areas is not valid as a fish of length 20 cm could be  
290 of age 1 or age 2 in Skagerak area. see Figure 3. This trend can be seen over several years of IBTS time  
291 series, which further discredit the assumption of constant ALK over large areas in the North sea. Figure S2  
292 in Supplementary Materials S7.1 illustrates that age given length of 20 cm long cod varies over large areas  
293 in years 2014-2018, while FigureS3 provides more illustrative examples of different length groups further

<sup>294</sup> discrediting the assumption of constant ALK over large areas in the North Sea.

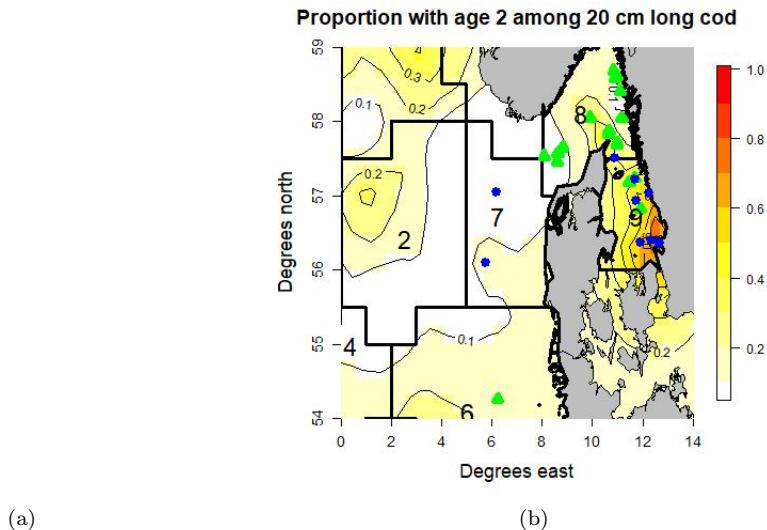
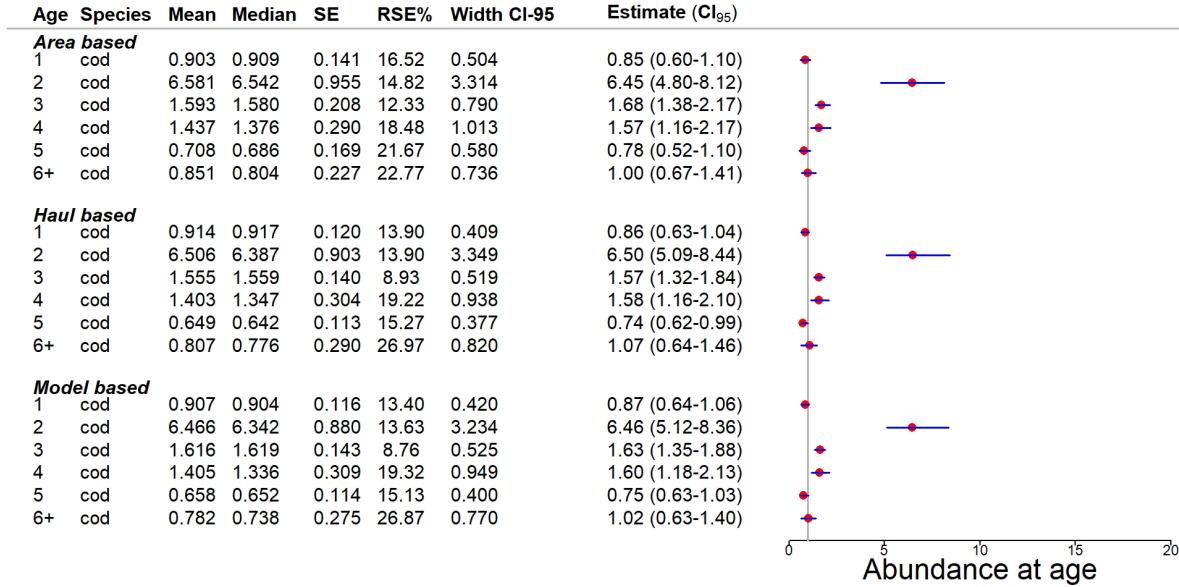


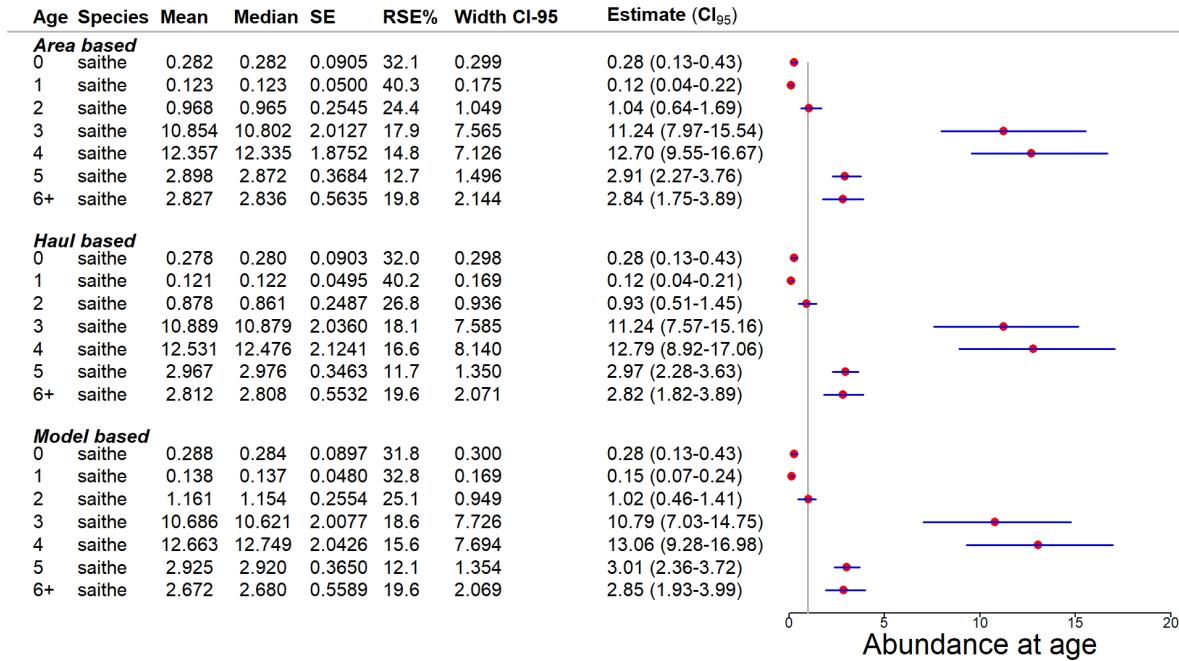
Figure 3: Estimated proportion of age 1 and 2 year-old cod of length 20 cm long in Skagerak. The green triangles ( $\blacktriangle$ ) and blue circles ( $\bullet$ ) are observations of one and two year old cod, respectively, which are in the length interval 19 cm to 21 cm.

Figures 4 gives estimates of indices of abundance for cod in 2018 Q1 and for saithe in 2017 Q3. Approximate 95% confidence intervals from the bias-corrected bootstrap method for 200 bootstrap replication are estimated from the three ALK methods I think we need to run 'production' run on larger number of iterations before interpreting too much. The stratified procedure described in 2.3.1 is used in the sampling process to estimate bootstrap confidence intervals. Figures 4 shows that the resulting indices of abundance for cod and saithe turned out to be similar for all ALKs. IBTS is a complex multistage survey design, and since the ALKs are estimated from cluster-correlated data the resulting effective sample for estimating age-composition of fish would be lower than the number of fish measured (ICES, 2013b). Hence, the ALKs are subject to large sampling errors. For example, the estimated percentage relative standard errors from the spatial ALKs for the plus group (6+) for cod are  $> 25\%$ , suggesting high sampling error in the ALKs. (Which parameter is tested here (age, length or something else)? Could the observation also be explained by high natural variation and the collapsing of potentially heterogeneous length and ages into one group? Also, it should be remembered that DATRAS ALK is constant. Aanes and Vølstad (2015) showed that in such cases, and where only the variability of length compositions are allowed for, the estimated age-distributions may appear to be more precise than they truly are since the ALK itself is subject to sampling errors, see

310 for example the estimated relative standard standard errors for ages 2, and the older fish (4, 5 and 6+) for  
 311 both species.



(a) Cod in year 2018 Q1



(b) Saithe in year 2017 Q3

Figure 4: Estimated confidence intervals (CI<sub>95</sub>) from bias-corrected bootstrap method for cod in year 2018 Q1 and saithe in year 2017 Q3. Estimated indices of abundance (Estimate), and its standard error (SE), percentage relative standard error (RSE%), bootstrap mean (Mean) and Median estimates and the width of the confidence interval (Width CI-95) are also given.

312 As regards to which spatial ALK method to adopt, it is difficult to identify a method that gives the  
 313 best performance over all age groups. While both methods seem to give reasonable estimates, the spatial  
 314 model based ALK generally gave shorter interval widths for both species (Figure 4). Furthermore, compared  
 315 with DATRAS ALK and the haul based ALK, the spatial model based ALK allows smooth functions of the  
 316 spatial effects predicting numbers-at-age. Figure 5 illustrates the estimated age compositions as a function  
 317 of length for a given haul in RFA 1. The haul selected is the haul with the most number of observed ages  
 318 of cod in 2018 Q1. Notice that the the model based ALK is smooth, while the DATRAS ALK and the haul  
 319 based ALK are not. This is an important advantage of the model based ALK, and it is surprising that it  
 320 did not result in a larger difference in the estimated index of abundance as shown Figure 4. An intuitive  
 321 reason for this is presumably because there are enough observed ages per length group for the haul based  
 322 ALK to be representative. But, there are some limitations of the spatial model based ALK. For instance,  
 323 the uncertainty of relative abundance from the spatial model based ALK is calculated using bootstrapping,  
 324 as approximation of the joint distribution of the regression coefficient and spatial effect, in some cases, fails  
 325 to account for the negative correlations between ages. Also, estimating relative abundance at age and its  
 326 precision from the spatial ALK model can be computationally intensive.

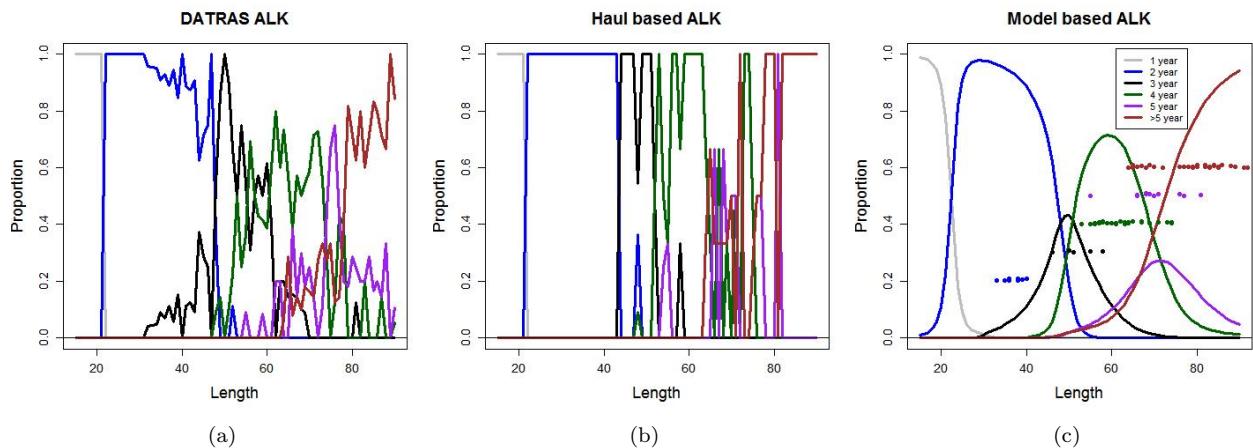


Figure 5: Estimated age compositions of cod as a function of length in a given haul in RFA 1 using a) DATRAS ALK, b) haul based ALK and c) model based ALK. Note that explanation of the colours are only given in c). Each coloured point in c) defines an observed cod with the corresponding length and age in the haul. The haul selected is the haul with most observed ages of cod in 2018 Q1.

327 We also demonstrate the implications of using DATRAS bootstrap procedure for estimating the uncer-

328 tainty around indices of abundance (see Figure S4 in Supplementary Materials S7.2). Compared with the  
329 stratified bootstrap procedure, DATRAS bootstrap procedure gives an overestimation of the uncertainty for  
330 all age groups, suggesting that it is highly relevant to account for the variation in the data over large areas.

331 ***3.2 Alternative sampling procedure for North Sea Cod and Saithe***

Table 1: Comparison of estimated uncertainty from proposed sampling procedures with estimated uncertainty from original data

age	mCPUE	SE.resampled	SE	RSE.resampled	RSE
<b>1 cm (85.9% of otoliths)</b>					
1	0.86	0.122	0.119	14.179	13.895
2	6.50	1.002	0.902	15.426	13.897
3	1.57	0.180	0.140	11.518	8.934
4	1.58	0.294	0.304	18.611	19.223
5	0.74	0.122	0.113	16.560	15.269
6+	1.07	0.275	0.289	25.623	26.968
<b>3 cm (62.1% of otoliths)</b>					
1	0.86	0.133	0.119	15.478	13.895
2	6.50	1.059	0.902	16.311	13.897
3	1.57	0.192	0.140	12.279	8.934
4	1.58	0.298	0.304	18.714	19.223
5	0.74	0.132	0.113	17.767	15.269
6+	1.07	0.311	0.289	29.178	26.968
<b>5 cm (51.0% of otoliths)</b>					
1	0.86	0.129	0.119	15.126	13.895
2	6.50	0.995	0.902	15.338	13.897
3	1.57	0.163	0.140	10.381	8.934
4	1.58	0.320	0.304	20.095	19.223
5	0.74	0.208	0.113	27.857	15.269
6+	1.07	0.289	0.289	27.193	26.968

332 In this section we investigate the effect of sampling fewer otoliths on the estimated indices of abundance  
 333 for the North Sea IBTS saithe and cod. **We use the spatial ALK model based approach, although**  
 334 **the haul based could also be used (see Supplementary Materials.....).** The removal procedure for  
 335 otolith sampling described in Section ?? is applied to data in year 2018 Q1 for cod and year 2017 Q3 for  
 336 saithe. We sample one pair of otoliths per length group described in Section ??: 1 cm, 2 cm, 3 cm, 4 cm,  
 337 5 cm, 6 cm or 7 cm. Recall that prior to 2018 the standardized IBTS sampling procedure was 8 pairs of  
 338 otoliths per length group but some nations such as Norway and Netherlands sampled one pair of otoliths  
 339 per length group from every haul. Although the revised standardized IBTS sampling procedure is one pair  
 340 of otolith per 1 cm length group for standard round fish as of year 2018 Q1, except for haddock and Norway  
 341 Pout where 2 otoliths per cm is to be sampled, some nations (Scotland and Sweden) continue to sample  
 342 more than one pair of otoliths, particularly for older age groups (see Table S2.1 in Supplementary Materials  
 343 S2).  
 344 Figure 6 gives the percentage relative standard error of estimated indices of abundance and mean square

345 error for cod and saithe from the seven different sampling procedures described above. Estimates are com-  
346 puted from 1000 simulations and 1000 bootstrap replication A total of 1600 pairs of otoliths were sampled  
347 for cod in year 2018 Q1, while 2163 pairs of otoliths were sampled for saithe in year 2017 Q3 (see Table  
348 S6.2 in Supplementary Materials S6). The proportion of otoliths removed for cod from each of the sam-  
349 pling procedures stated above is: 14.4%, 28.6%, 38.4%, 44.5%, 49.3%, 52.6% or 55.6%, respectively, while  
350 for saithe the following proportions of otoliths are removed: 27.1%, 48.9%, 59.5%, 65.6%, 69.8%, 73.1% or  
351 75.2%, respectively. Notice that 14% of the cod data in year 2018 Q1 is removed for the sampling procedure  
352 of a pair of otoliths per 1 cm length group. This should be 0% if all nations followed the revised standardized  
353 IBTS sampling procedure of year 2018 Q1.

354

355 **Tables S7.2 and S7.3 in Supplementary Materials S7.3 give results of the estimated indices**  
356 **of abundance and approximate 95% bias-corrected bootstrap confidence intervals**  
357 **discuss graph**

358

359 • **We discuss and include these in explanations below**

360 • Accuracy of estimates of reduced data compared with estimates from full data

361 • Precision in estimates is measured by standard error (SE) and relative standard error (RSE)

362 • accuracy is measured by root mean square error (RMSE) =  $\sqrt{SE^2 + (\text{bias})^2}$ . Measures how close, on  
363 average, a fitted line is to the data points (measure of goodness of fit). One can compare the RMSE to  
364 observed variation in measurements of a typical point (**the two should be similar for a reasonable**  
365 **fit**). Can we use this even though we do not have a "true value", which we would never know from  
366 large survey data and since we did not simulate synthetic data? Can we consider  $\hat{\lambda}_a$  as a "true value"?

367 The nonparametric bias-corrected bootstrap method is adopted for estimating confidence intervals of  
368 indices of abundance, and although this method has the advantage of correcting for the bias and skew of  
369 the sampling distribution of the data; accounting for some of the variability in the sampling distribution of  
370 the CPUE; and does not assume any distribution for the data, there are some limitations of the bootstrap

approach. The most important limitation is the assumption that the distribution of the data represented by the sample is a reasonable estimate of the population function from which the data are sampled. If this assumption is violated the random sampling performed in the bootstrap procedure may add another level of sampling error, resulting in invalid statistical estimations (Haukoos and Lewis, 2005). As discussed in Section 1.1 the selection of the trawling locations for IBTS surveys is semi-random where cruise leaders selects "clear" tow locations or "blind" tow locations if no clear tow exists by checking the proposed trawl track for hazardous seabed obstructions with acoustic methods. More recently, selection of tow locations is based on pre-proposed valid tow locations with start and end positions executed in the period 2000-2018. Hence, the lack of a fully randomized sampling process has the potential to result in biased estimates of parameters and their uncertainty. Additionally, prior to 2013, all nations were sampling 8 pairs of otoliths per 1 cm length group for our focal species (Table S2.1 in Supplementary Materials S2), and these samples could be acquired from, for example the first haul (or first few trawl hauls), resulting in an unrepresentative sample of the population. From 2013, some nations adopted the current sampling procedure outlined by ICES for IBTS 2018 surveys of 1 pair of otolith per 1 cm length group from each haul, while other nations continued with sampling 8 pairs of otoliths per 1 cm length group. So, bias was still introduced via the sampling procedure. Another limitation of the bootstrap is the smaller the original sample the less likely it is to represent the entire population, thus the more difficult it becomes to compute valid confidence intervals.

Note that the bootstrap relies heavily on the tails of the estimated sampling distribution when computing

389

390       **these results in the graph are from the haul based ALK procedure. The model based**  
391       **ALK procedure gave an error, when it's working those will be here and haul based will go in**  
392       **supplementary materials**

393

## 4 DISCUSSION

394 In this research we have determined optimal sampling efforts of otoliths for target species of the North Sea  
395 International Bottom Trawl Survey (IBTS). This was achieved by testing different sampling procedures that  
396 mimic the real data collection procedure but with a reduced number of otoliths. The estimated indices of

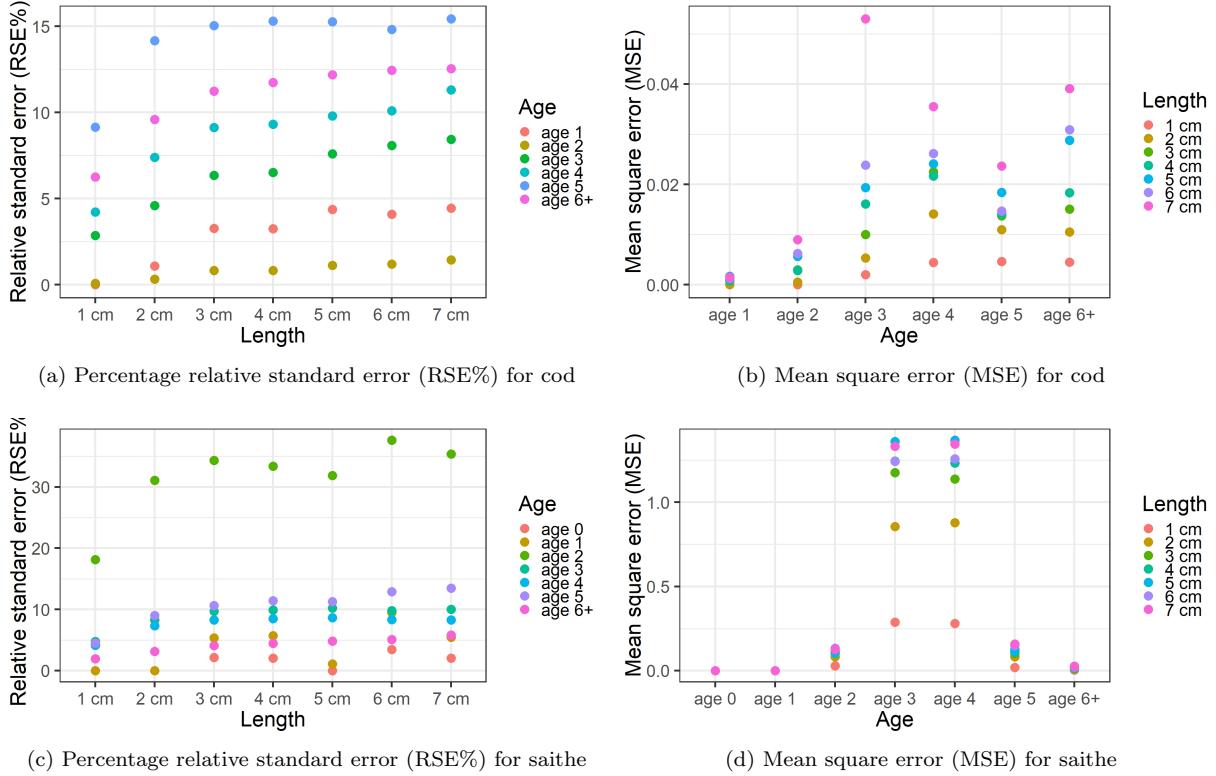


Figure 6: Percentage relative standard error (RSE%) and mean square error (MSE) for age given seven length group sampling procedures of otolith collection for cod in year 2018 Q1 and saithe in year 2017 Q3.

abundance and their estimated uncertainty were investigated to determine if there is any real change in the precision of the estimates. Abundance indices were estimated using age-length keys (ALKs). The database for trawl surveys (DATRAS) manned by ICES includes an ALK that uses the raw proportions of age given length assuming constant age-length compositions over relatively large areas. We have developed two spatial ALK methods to estimate abundance indices and their variance that accounts for spatial variation in the data: 1) a haul based ALK that produces an ALK for each trawl haul, and which uses the raw proportions of age given length, and 2) a spatial ALK model that uses logits for modelling the age distribution in catch data from the length-stratified subsamples. Several studies have used spatial ALK modelling for estimating abundance indices of the North Sea stocks used in assessments (Berg and Kristensen, 2012; Berg et al., 2014; Gerritsen et al., 2006). These studies used continuous ratio logits with General Linear Model (GLM) or General Additive Models (GAMs) to model the spatial effects and found large spatio-temporal variability of the ALK and relative abundance at age. We proposed to use Gaussian Random Field Theory to model the

409 spatial effects as a smooth surface to estimate age-at-length and relative abundance for the IBTS data. The  
410 spatial model based ALK and the design based spatial ALK (haul based) gave similar estimates as DATRAS  
411 estimator for relative abundance at age but the spatial ALK estimators gained better precision.

412 The spatial ALK model based estimator appears to be a useful tool to detect significant differences  
413 between ALKs over large areas, although estimation of the uncertainty in the ALK from the joint precision  
414 matrix is problematic. Including the uncertainty of the ALK in the model requires an approximation of the  
415 joint distribution of the regression coefficient and the spatial effect, but this approximation is only as good as  
416 the quality of the data in a given year and quarter. For instance, the approximation of the ALK can predict  
417 juvenile ages given longer lengths, which goes against the natural biology. This occurs presumably because  
418 the approximation fails to account for the negative correlation structures between ages. So the uncertainty  
419 in the relative abundance was, therefore, calculated using bootstrapping as done by Berg and Kristensen  
420 (2012); Berg et al. (2014). In future, the model might be expanded to include the probability of recording  
421 inaccurate age-at-length data, so that uncertainty in the ALK could be estimated using the joint precision  
422 matrix. The model might also be expanded to include covariates such as trawl hauls to capture any haul  
423 variation, for example a trawl haul may "hit" a school of fish of a certain age.

424 With regards to how many otoliths to sample per length group, the evidence is clear that .....

425

426 **discuss DATRAS and Haul based ALK and recommended optimum sampling level of**  
427 **otoliths per length group**

## References

- 429 Aanes, S. and Vølstad, J. H. (2015). Efficient statistical estimators and sampling strategies for estimating  
 430 the age composition of fish. *Canadian journal of fisheries and aquatic sciences*, 72(6):938–953.
- 431 Berg, C. W. and Kristensen, K. (2012). Spatial age-length key modelling using continuation ratio logits.  
 432 *Fisheries Research*, 129:119–126.
- 433 Berg, C. W., Nielsen, A., and Kristensen, K. (2014). Evaluation of alternative age-based methods for  
 434 estimating relative abundance from survey data in relation to assessment models. *Fisheries Research*,  
 435 151:91–99.
- 436 Breivik, O. N. (2018). spatialalk. <https://github.com/OlavNikolaiBreivik/IBTSspatialALK>.
- 437 Carpenter, J. and Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? a practical guide  
 438 for medical statisticians. *Statistics in medicine*, 19(9):1141–1164.
- 439 Ehrhardt, N. M. and Legault, C. M. (1997). The role of uncertainty in fish stock assessment and management:  
 440 a case study of the spanish mackerel, scomberomorus maculatus, in the us gulf of mexico. *Fisheries research*,  
 441 29(2):145–158.
- 442 Fridriksson, A. (1934). On the calculation of age-distribution within a stock of cod by means of relatively few  
 443 age-determinations as a key to measurements on a large scale. *Rapports Et Proces-Verbaux Des Reunions*,  
 444 *Conseil International Pour l'Exploration De La Mer*, 86:1–5.
- 445 Gerritsen, H. D., McGrath, D., and Lordan, C. (2006). A simple method for comparing age-length keys  
 446 reveals significant regional differences within a single stock of haddock (*melanogrammus aeglefinus*). *ICES*  
 447 *Journal of Marine Science*, 63(6):1096–1100.
- 448 Haukoos, J. S. and Lewis, R. J. (2005). Advanced statistics: bootstrapping confidence intervals for statistics  
 449 with difficult distributions. *Academic emergency medicine*, 12(4):360–365.
- 450 Hirst, D., Storvik, G., Rognebakke, H., Aldrin, M., Aanes, S., and Vølstad, J. H. (2012). A bayesian  
 451 modelling framework for the estimation of catch-at-age of commercially harvested fish species. *Canadian*  
 452 *journal of fisheries and aquatic sciences*, 69(12):2064–2076.

- 453 ICES (2006). Report of the workshop on implementation in datras of confidence limits estimation of abundance indices from bottom trawl survey data. International Council for the Exploration of the Sea, ICES
- 454 DATRAS REPORT.
- 456 ICES (2012). Manual for the international bottom trawl surveys, revision viii. series of ices survey protocols.
- 457 International Council for the Exploration of the Sea, SISP 1-IBTS VIII.
- 458 ICES (2013a). Ns-ibts indices calculation procedure. datras procedure document. International Council for
- 459 the Exploration of the Sea, 1.1 NS-IBST indices-2013.
- 460 ICES (2013b). Report of the third workshop on practical implementation of statistical sound catch sampling
- 461 programmes. International Council for the Exploration of the Sea, WKPICS3.
- 462 ICES (2015). Manual for the international bottom trawl surveys, revision ix. series of ices survey protocols.
- 463 International Council for the Exploration of the Sea, SISP 10-IBTS IX.
- 464 ICES (2016a). Ices advice on fishing opportunities, catch, and effort greater north sea and celtic seas
- 465 ecoregions. International Council for the Exploration of the Sea, Book 6.
- 466 ICES (2016b). Report of the benchmark workshop on north sea stocks (wknsea), 14–18 march, 2016.
- 467 International Council for the Exploration of the Sea, ICES CM 2016/ACOM:37.
- 468 ICES (2018a). Ices fishmap: Atlas of the north sea fish, including fact sheets of key species and distribution
- 469 maps. International Council for the Exploration of the Sea (ICES FishMap website), 2018, 1.
- 470 ICES (2018b). Manual for the international bottom trawl surveys, revision x. series of ices survey protocols.
- 471 International Council for the Exploration of the Sea, SISP X-IBTS X.
- 472 ICES, D. (2018c). Datras.
- 473 Karlsson, A. (2009). Bootstrap methods for bias correction and confidence interval estimation for nonlinear
- 474 quantile regression of longitudinal data. Journal of Statistical Computation and Simulation, 79(10):1205–
- 475 1218.

- 476 Kimura, D. K. (1977). Statistical assessment of the age-length key. Journal of the Fisheries Board of Canada,  
477 34(3):317–324.
- 478 Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. (2015). Tmb: automatic differentiation  
479 and laplace approximation. arXiv preprint arXiv:1509.00660.
- 480 Lehtonen, R. and Pahkinen, E. (2004). Practical methods for design and analysis of complex surveys. John  
481 Wiley & Sons.
- 482 Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between Gaussian fields  
483 and Gaussian Markov random fields: the stochastic partial differential equation approach.  
484 Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73(4):423–498.
- 485 Ludwig, D. and Walters, C. J. (1981). Measurement errors and uncertainty in parameter estimates for stock  
486 and recruitment. Canadian Journal of Fisheries and Aquatic Sciences, 38(6):711–720.
- 487 Mashreghi, Z., Haziza, D., Léger, C., et al. (2016). A survey of bootstrap methods in finite population  
488 sampling. Statistics Surveys, 10:1–52.
- 489 Puth, M.-T., Neuhäuser, M., and Ruxton, G. D. (2015). On the variety of methods for calculating confidence  
490 intervals by bootstrapping. Journal of Animal Ecology, 84(4):892–897.
- 491 Rindorf, A. and Lewy, P. (2001). Analyses of length and age distributions using continuation-ratio logits.  
492 Canadian Journal of Fisheries and Aquatic Sciences, 58(6):1141–1152.
- 493 Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference  
494 for latent Gaussian models by using integrated nested Laplace approximations.  
495 Journal of the Royal Statistical Society: Series B (Statistical Methodology), 71(2):319–392.
- 496 Stein, M. L. (2012). Interpolation of spatial data: some theory for kriging. Springer Science & Business  
497 Media.
- 498 Walters, C. J. and Ludwig, D. (1981). Effects of measurement errors on the assessment of stock-recruitment  
499 relationships. Canadian Journal of Fisheries and Aquatic Sciences, 38(6):704–710.

- 500 Wood, S. (2015). Package mgcv. R package version, pages 1–7.
- 501 Wood, S. N. (2017). Generalized additive models: an introduction with R. CRC press.

502 **Supplemental Materials: Optimizing sampling effort of the North**  
503 **Sea International Bottom Trawl Survey.**

504 **S1 Areas fished by different countries in the North Sea IBTS**

505 Typically, two different countries fish each rectangle so that at least two trawl hauls are made per rectangle,  
506 but intensified sampling is carried out in the following areas: at least 3 hauls per rectangle are taken in  
507 statistical rectangles 31F1, 31F2, 32F1, 33F4, 34F2, 34F3, 34F4, 35F3, 35F4; while six or more hauls per  
508 rectangle are taken in statistical rectangles 30F1, 32F2, 32F3, 33F2, 33F3 (ICES 1999). The Skagerrak  
509 and Kattegat is fished solely by Sweden, who sample more than once in every rectangle while the west of  
510 Shetland (in Q1 and Q3) and inshore areas (Q3) is fished solely by Scotland. The edge of the Norwegian  
511 Trench is fished solely by Norway, but inshore areas near Denmark is fished by Denmark. The southern  
512 North Sea is fished by Denmark, Germany and England. France, typically, is the only country that surveys  
513 the western English Channel. Areas are surveyed by a single country because of the large proportion of  
514 untrawalable area (and subsequent gear damage issues experienced by other nations) for efficient logistical  
515 purposes. Table S1.1 gives the countries and research vessels participating the North Sea IBTS.

Table S1.1: Survey countries, vessel name, and period research vessels participating in first quarter (Q1) and third quarter (Q3) during 1997-2017.

Country	First Quarter (Q1)		Third Quarter (Q3)	
	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

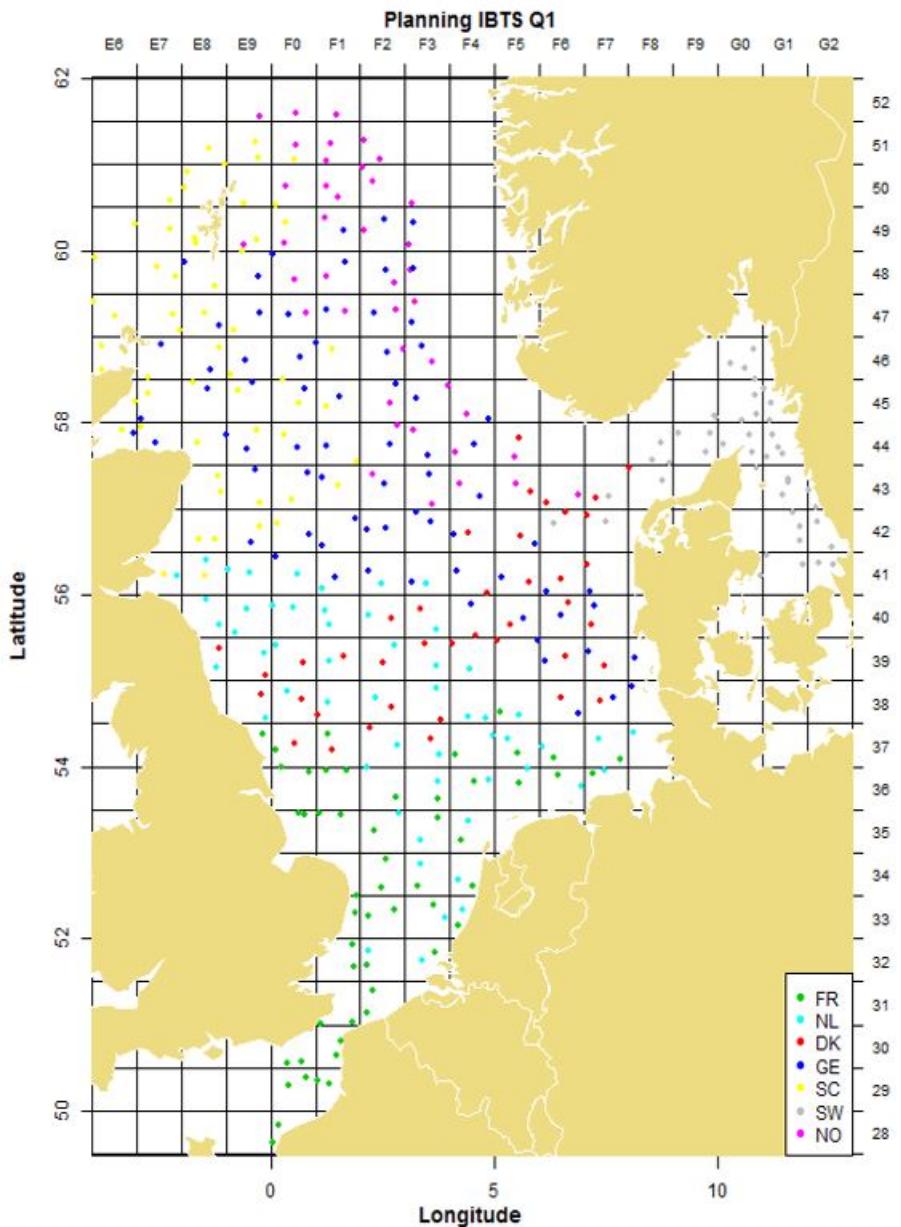


Figure S1: North Sea IBTS 2018 sampling program for nations. Seven nations carry out sampling: France (FR), Netherlands (NL), Denmark (DN), Germany (GE), Scotland (SC), Sweden (SW) and Norway (NO).

517

## S2 Otolith sampling per fish species

- 518 From 1991-2017, most countries conducted quota sampling of otoliths per length group in a RFA. But from  
 519 2013 Norway has been sampling one otolith per length class from each trawl haul (to 0.1cm below for shellfish,

520 to 0.5cm below for herring and sprat and to 1cm below for all other species). From the first quarter in 2018  
 521 all countries are required to sample one otolith per length class per trawl haul. Table S2.1 gives the minimum  
 522 sampling levels of otoliths for the target species. However, for the smallest size groups, that presumably  
 523 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 S2.1: Minimum sampling levels of otoliths by species for RFA or per trawl haul.

Period	Species	Minimum sampling levels of otoliths per length class
<b>1991-2017</b>		
		<b>Number of otoliths per length class in a RFA</b>
	herring	8 otoliths 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 $\geq 11.0$ 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 otoliths per 1 cm length class
	Norway pout	8 otoliths per 1 cm length class
	saithe	8 otoliths per 1 cm length class
	All target species	From 2013 Norway and Scotland, and Netherlands from 2016 have 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).
<b>2018</b>		
		<b>Number of otoliths per length class per trawl haul</b>
	herring	1 otolith per $\frac{1}{2}$ cm group
	sprat	1 otolith per $\frac{1}{2}$ cm length class 8.0 – 11.0 cm
		1 otolith per $\frac{1}{2}$ cm length class $\geq 11.0$ cm
	mackerel	1 otolith per 1 cm length class
	cod	1 otolith per 1 cm length class
	haddock	2 otoliths per 5 cm length class 11 – 15, 16 – 20, 21 – 25, 26 – 30 cm
	Norway pout	2 otoliths per 5 cm length class 5 – 10, 11 – 15 cm 2 otoliths per 1 cm length class $> 15$ cm
	saithe	1 otolith per 1 cm length class
	plaice	1 otolith per 1 cm length class

### S3 Weightings of Statistical Rectangles

526 The weightings of the some statistical rectangles are allotted to species such as sprat, saithe and herring by  
 527 depth strata between 10m -200m and for RFA 8 and 9 water depths between 10m-250m. Table S3.1 gives  
 528 these weights, which are used in the analysis of the saithe data (ICES, 2013a).

Table S3.1: Weights used for *Pollachius virens* in equation (2.3).

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

530

## S4 Imputing missing ages

531 With the current IBTS sampling procedure, *one* age reading is collected from every cm group in every haul  
 532 for cod and saithe. However, missing ages do occur, and in this section we describe how missing ages are  
 533 accounted for. Note that missing ages are automatically accounted for with the model based approach.  
 534 However, for the area based ALK and the haul based ALK, some adhoc procedure needs to be followed when  
 535 age is missing. We want to highlight that missing ages occur relatively seldom. E.g. in 1st. quarter in year  
 536 2018 and with length group width equal 1cm, 0.9% of the cod miss age with respect to the area based ALK.  
 537 In the same example, 1.4% of the cod miss age with respect to the haul based ALK.

538 ***S4.1 Area based ALK***

539 Assume there do not exist an age reading of a fish of length  $l$  in a RFA. Just as in (ICES, 2013a), the  
 540 following procedure is done for the area based ALK:

- 541 • If  $l$  is between the minimum length and maximum length, the age is set to be equal the ALK to the  
 542 closest length group with observed ages in the RFA. In cases where there are two equally close length  
 543 groups with observed ages, the average of those two ALKs is used.
- 544 • If  $l$  is smaller then the smallest measured fish in the RFA, the age is set to the minimum age.
- 545 • If  $l$  is larger or equal the maximum length, the age is set to the maximum age.

546 ***S4.2 Haul based ALK***

547 Assume there do not exist an age reading of a fish of length  $l$  in a haul. For the haul based procedure, the  
 548 following is done in sequential order:

- 549 • If there exist an age reading of that length group closer than 60 nautical mile in the same RFA, the  
 550 ALK from the closest haul with such an age reading is used.
- 551 • If there exist a fish with length in the interval  $l \pm 1\text{cm}$  with age information in the same haul, that  
 552 observed age is used.

- 553     • If the two first step did not produce an ALK for  $l$ , there exist little information close in space and  
 554         length, and the area based ALK is used.

555           **S5 Nonparametric Bootstrap Sampling procedure**

556     Nonparametric bootstrapping is attractive as it makes no distributional assumption, and is suitable for  
 557         estimating confidence interval for indices of abundance. Suppose we have a vector  $\mathbf{x}$  of  $m$  independent obser-  
 558         vations, and we are interested in estimating a parameter  $\hat{\theta}(\mathbf{x})$  and its variance. The general nonparametric  
 559         bootstrap algorithm is as follows:

560         1. Sample  $m$  observations randomly with replacement from  $\mathbf{x}$  to obtain a bootstrap data set, denoted by

561              $\mathbf{x}^*$ .

562         2. Calculate the bootstrap version of the statistic of interest,  $\theta^* = \hat{\theta}(\mathbf{x}^*)$ .

563         3. Repeat steps 1 and 2 a large number of times, say  $B$ , to obtain an estimate of the bootstrap distribution

564         4. calculate the average of the bootstrapped statistics,  $\sum_{b=1}^B \theta^*_{(b)} / B$

565         5. compute the variance of the estimator  $\hat{\theta}(\mathbf{x})$  through the variance of the set  $\theta^*_{(b)}$ ,  $b = 1, 2, \dots, B$ , given

566         by

$$\frac{\sum_{b=1}^B (\theta^*_{(b)} - \theta^*_{(\cdot)})^2}{(B - 1)} \quad (\text{S5.1})$$

567         where  $\theta^*_{(\cdot)} = \sum_{b=1}^B \theta^*_{(b)} / B$ .

568     The Bias-Corrected method assumes that there is a monotonic increasing function and the estimator  $\hat{\lambda}_a$  has  
 569         a monotonic increasing function  $f()$  such that the transformed values  $f(\hat{\lambda}_a)$  are normally distributed with  
 570         mean  $f(\lambda_a) - z_0$  and standard deviation one, where  $z_0$  are the standard normal limits (Puth et al., 2015;  
 571         Karlsson, 2009). Now, let  $P^* \left( \hat{\theta}(\mathbf{x}^*) \leq \hat{\theta}(\mathbf{x}) \right)$  denote the proportion of  $\hat{\theta}(\mathbf{x}^*)$ 's in the bootstrap sample that  
 572         have a value lower than the value of the parameter estimate  $\hat{\theta}(\mathbf{x})$ , and let  $z_0$  be defined as

$$z_0 = \Phi^{-1} \left\{ P^* \left( \hat{\theta}(\mathbf{x}^*) \leq \hat{\theta}(\mathbf{x}) \right) \right\}, \quad (\text{S5.2})$$

<sup>573</sup> where  $\Phi$  denotes the cumulative distribution function of the standard normal distribution. Also let  $\tilde{\alpha}_1$  and

<sup>574</sup>  $\tilde{\alpha}_2$  be defined as

$$\tilde{\alpha}_1 = \Phi(2z_0 + z_\alpha), \quad (\text{S5.3})$$

<sup>575</sup> and

$$\tilde{\alpha}_2 = \Phi(2z_0 + z_{1-\alpha}), \quad (\text{S5.4})$$

<sup>576</sup> respectively. A  $100(1 - 2\alpha)$  percent confidence interval for  $\theta(\mathbf{x})$  is then given by

$$\hat{\theta}_{(\tilde{\alpha}_1(B+1))}(\mathbf{x}^*) \leq \hat{\theta}(\mathbf{x}) \leq \hat{\theta}_{((\tilde{\alpha}_2-1)(B+1))}(\mathbf{x}^*). \quad (\text{S5.5})$$

## <sup>577</sup> S6 IBTS data set for cod and saithe

Table S6.1: Age and length data for saithe and cod . Data collected in the first quarter (Q1) has no age 0 group but this is collected in quarter 3 (Q3) surveys.

Species	Years									
	2014		2015		2016		2017		2018	
	Q1	Q3	Q1	Q3	Q1	Q3	Q1	Q3	Q1	Q3
Trawls	352	333	387	367	370	389	385	345	372	365
Mean time of hauls (in minutes)	29.20	28.91	28.55	23.44	28.40	22.47	29.02	29.37	29.26	29.13
<b>cod</b>										
Age range (in years)	1 - 9	0 - 10	1 - 9	0 - 8	1 - 10	0 - 10	1 - 11	0 - 11	1 - 12	0 - 11
Length range (in cm)	8 - 113	NA	9 - 113	7 - 111	11 - 109	5 - 110	7 - 115	6 - 112	11 - 114	5 - 107
Total otoliths	1903	2399	2895	2113	2046	1804	2501	2230	1600	1456
<b>saithe</b>										
Age range (in years)	1 - 16	0 - 15	1 - 16	0 - 19	1 - 17	0 - 19	1 - 15	0 - 11	1 - 12	0 - 13
Length range (in cm)	15 - 105	8 - 108	13 - 110	NA	16 - 107	12 - 115	15 - 108	6 - 112	11 - 114	7 - 109
Total otoliths	584	1182	600	1526	581	1631	1083	2163	822	1085

Table S6.2: Summary of North Sea IBTS cod and saithe (in parentheses) data for third quarter in year 2017 and first quarter in year 2018.

Data	Description
Trawl hauls	Total of 372 trawl hauls in year 2018 Q1; 238 (83) with length and 230 (81) with age information. In 2017 Q3, a total of 345 trawl hauls were taken; 238 (129) with length and 237 (128) with age information.
Age	The age varied between 1 (1) to 12 (18) years in year 2018 Q1 and 0 (0) to 11 (17) in year 2017 Q3.
Length	Length information in cm varied between 11 (13) to 114 (106) cm in year 2018 Q1 and between 6 (10) to 112 (109) cm in year 2017 Q3.
Date	Date of catch in year 2018 Q1 varied between 15.01.2018 to 28.02.2018 and in year 2017 Q3 between 18.07.2017 to 31.08.2018
Duration of haul	Mean duration is 29.37 minutes, with 30 minutes as 83% coverage interval in year 2018 Q1; and in 2017 Q3 with mean duration of 29.26 minutes with 30 minutes as 88% coverage .
Total count for all ages	1600 (822) in year 2018 Q1 and 2330 (2163) 2017 Q3.

578

## S7 Analysis of real data

### 579 *S7.1 Proportion of age given length of cod*

580 Figure S2 shows that age-length structure of cod varies over several years: the first quarter (Q1) in years  
 581 2014-2017. This indicates that the assumption of constant ALK over large areas proposed in DATRAS is  
 582 invalid. The plots in Figure S3, which shows the proportion of age given different lengths of cod in 2018 Q1  
 583 further discredit this assumption as a fish of length 30 cm can be of age 2 (green triangle ▲) or age 3 (blue  
 584 circle ●) in the northern North Sea; see Figure S3 (a).

Table S6.3: Age and length data for saithe in year 2017 Q3 and cod in year 2018 Q1. Data collected in the first quarter (Q1) has no age 0 group but this is collected in quarter 3 (Q3) surveys.

Age	saithe in year 2017 Q3						cod in year 2018 Q1					
	Numbers aged	L <sub>min</sub>	L <sub>max</sub>	L <sub>mean</sub>	Sd(L)	CV(L)	Numbers aged	L <sub>min</sub>	L <sub>max</sub>	L <sub>mean</sub>	Sd(L)	CV(L)
0	21	10	14	12.143	1.195	0.098						
1	26	23	32	27.654	2.297	0.083	149	11	30	18.407	3.693	0.201
2	65	27	47	38.077	3.337	0.088	814	17	53	33.180	6.290	0.190
3	531	34	56	42.041	3.785	0.090	222	30	81	50.654	10.185	0.202
4	767	35	73	48.261	4.521	0.094	189	43	92	64.479	8.399	0.130
5	334	46	78	56.876	6.105	0.107	102	54	96	76.627	9.594	0.125
6	159	50	91	66.025	7.137	0.108	84	54	100	80.871	9.456	0.117
7	127	57	93	73.976	7.163	0.097	28	58	110	84.086	11.308	0.134
8	69	63	94	77.725	7.010	0.090	4	80	94	85.500	6.455	0.075
9	18	64	97	85.333	7.499	0.088	5	66	96	83.400	11.305	0.061
10	22	84	107	92.364	5.803	0.063	1	87	87	-	-	-
11	5	79	102	92.800	9.311	0.100	1	106	106	-	-	-
12	7	91	109	99.429	6.554	0.066						
13	5	94	104	98.800	4.550	0.046						
14	1	108	108	108	-	-						
15	1	105	105	108	-	-						
16	4	93	106	100.250	5.439	0.054						
17	1	109	109	109	-	-						

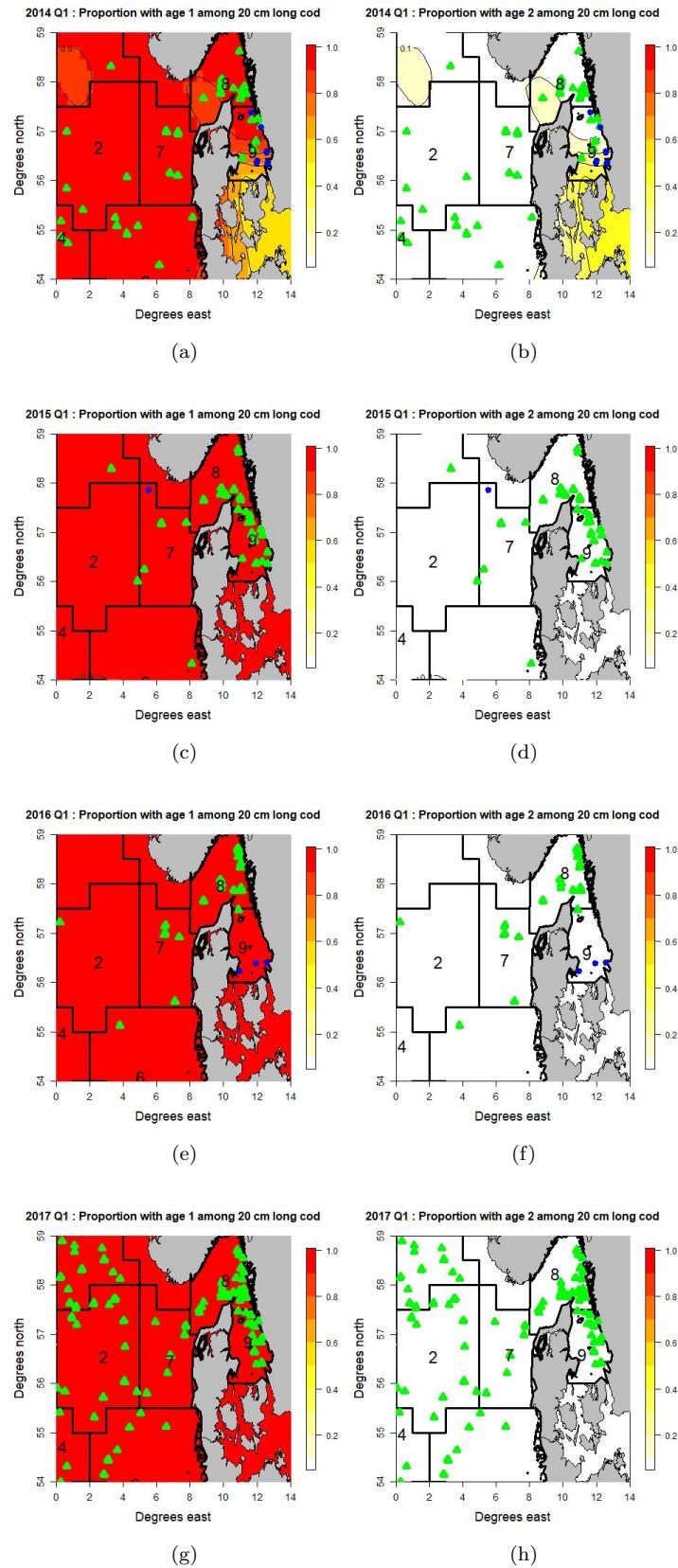
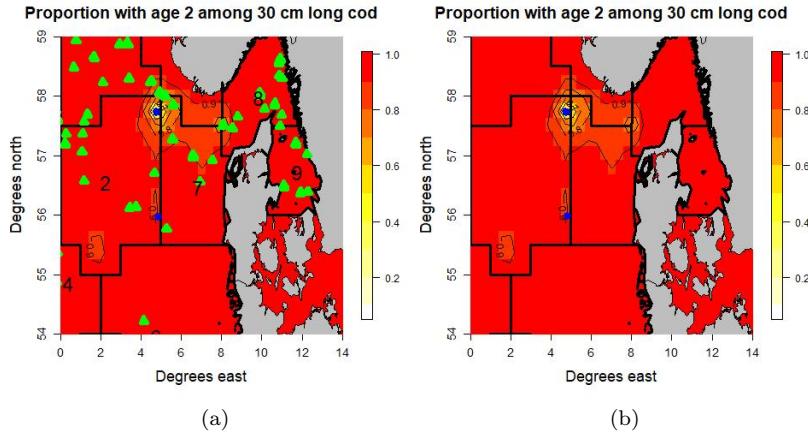
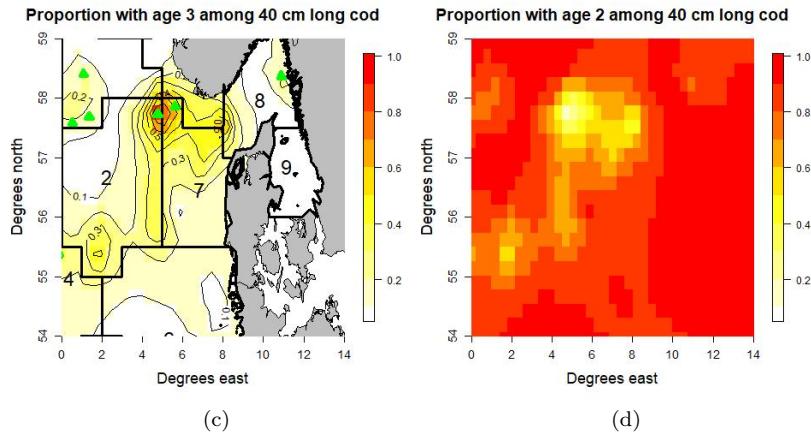


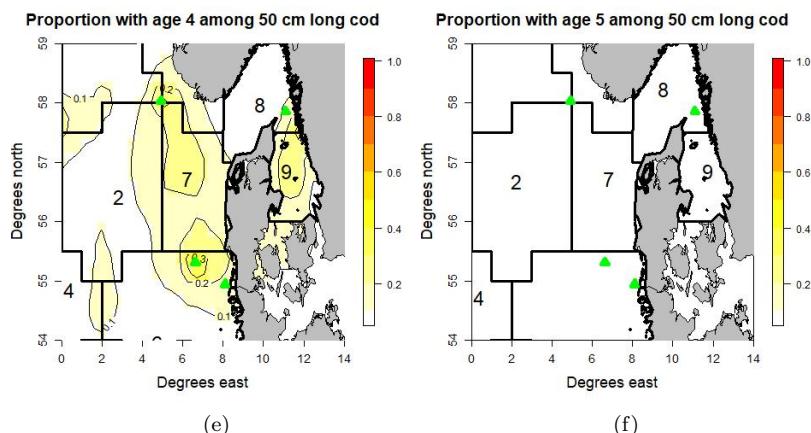
Figure S2: Estimated proportion of age 1 and 2 year-old cod of length 20 cm in the North Sea in the first quarter (Q1) in years 2014-2017. The green triangles ( $\blacktriangle$ ) are observations of 1-year olds and the blue circles ( $\bullet$ ) are observations of 2-year olds in the length interval 19 cm to 21 cm.



Estimated proportion of age 2 and 3 year-old cod of length 30 cm in the North Sea, where ▲ are observations of 2-year olds and ● are 3-year olds in the length interval 29 cm to 31 cm in 2018 Q1.



Estimated proportion of age 3 and 4 year-old cod of length 40 cm in North Sea, where ▲ are 3-year olds in the length interval 39 cm to 41 cm in 2018 Q1.



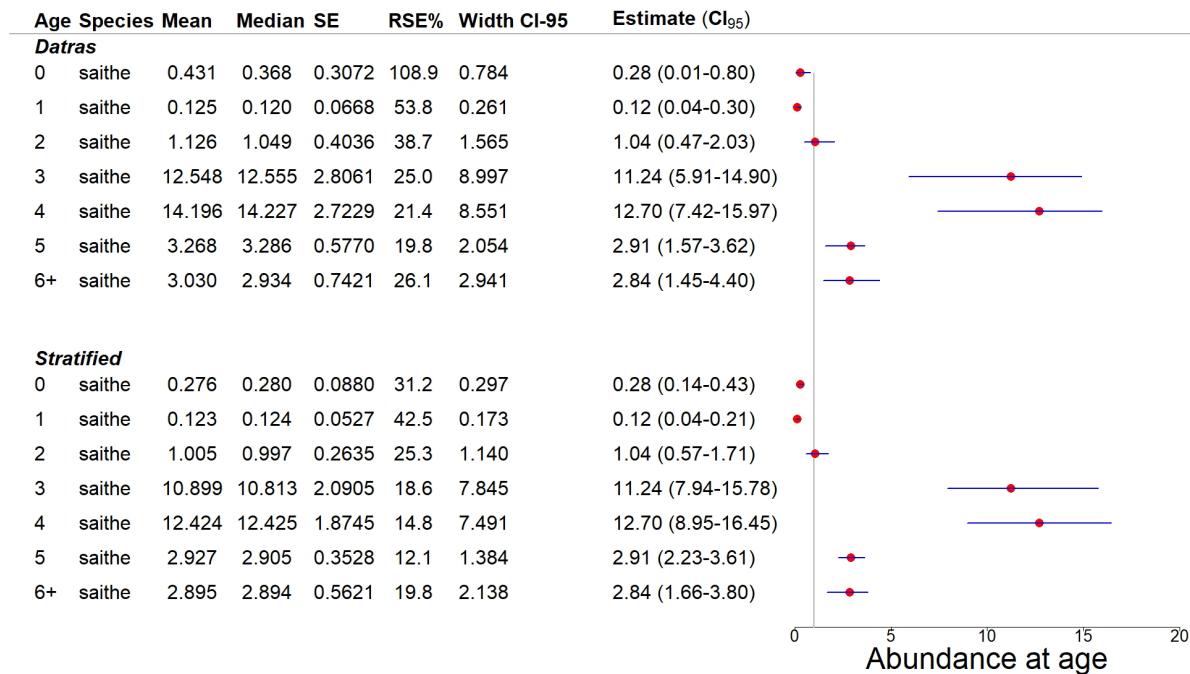
Estimated proportion of age 4 and 5 year-old cod of length 50 cm in Skagerak, where ▲ are 4-year olds in the length interval 49 cm to 51 cm in 2018 Q1.

Figure S3: Proportion of age given lengths 20 cm, 40 cm or 50 cm of cod in 2018 Q1 in the North Sea.

585 **S7.2 Estimates from DATRAS and Stratified bootstrap procedures**

586 The bootstrap procedure proposed by DATRAS lacks the potential to account for the spatial variation in the data.  
 587 The DATRAS bootstrap procedure ignores the fine-scale stratification in the sampling process, leading to an overesti-  
 588 mation of the uncertainty; and ignores the age-length data collected at the haul level, resulting in an underestimation  
 589 of the uncertainty. The results (FigureS4) shows an overestimation of the uncertainty for all age groups, suggesting  
 590 that it is relevant to account for the fine-scale stratification when resampling the data.

591



(a) Datras and Stratified bootstrap Procedures

Figure S4: Comparison of estimated confidence intervals (CI<sub>95</sub>) from DATRAS and stratified bootstrap procedures. The bias-corrected bootstrap method is used to give estimates for saithe in year 2017 Q3. Estimated indices of abundance (Estimate), and its standard error (SE), bootstrap mean (Mean), Median estimates, percentage relative standard error (RSE %) and width of confidence intervals are also given.

592 **S7.3 Estimates from different sampling procedures**

Table S7.1: Comparison of estimated uncertainty from proposed sampling procedures with estimated uncertainty from original data

age	mCPUE	SE.resampled	SE	RSE.resampled	RSE
<b>1 cm</b>					
1	0.86	0.122	0.119	14.179	13.895
2	6.50	1.002	0.902	15.426	13.897
3	1.57	0.180	0.140	11.518	8.934
4	1.58	0.294	0.304	18.611	19.223
5	0.74	0.122	0.113	16.560	15.269
6+	1.07	0.275	0.289	25.623	26.968
<b>2 cm</b>					
1	0.86	0.126	0.119	14.697	13.895
2	6.50	1.048	0.902	16.149	13.897
3	1.57	0.161	0.140	10.239	8.934
4	1.58	0.299	0.304	18.769	19.223
5	0.74	0.141	0.113	19.257	15.269
6+	1.07	0.263	0.289	24.643	26.968
<b>3 cm</b>					
1	0.86	0.133	0.119	15.478	13.895
2	6.50	1.059	0.902	16.311	13.897
3	1.57	0.192	0.140	12.279	8.934
4	1.58	0.298	0.304	18.714	19.223
5	0.74	0.132	0.113	17.767	15.269
6+	1.07	0.311	0.289	29.178	26.968
<b>4 cm</b>					
1	0.86	0.132	0.119	15.384	13.895
2	6.50	0.938	0.902	14.476	13.897
3	1.57	0.187	0.140	11.844	8.934
4	1.58	0.349	0.304	21.842	19.223
5	0.74	0.159	0.113	21.702	15.269
6+	1.07	0.280	0.289	26.335	26.968
<b>5 cm</b>					
1	0.86	0.129	0.119	15.126	13.895
2	6.50	0.995	0.902	15.338	13.897
3	1.57	0.163	0.140	10.381	8.934
4	1.58	0.320	0.304	20.095	19.223
5	0.74	0.208	0.113	27.857	15.269
6+	1.07	0.289	0.289	27.193	26.968

Table S7.2: Estimated abundance ( $\hat{\lambda}_a$ ) for cod from the original data in year 2018 Q1 compared with estimated abundance ( $\hat{\lambda}_a^*$ ) from the reduced data for different sampling procedures of length groups ( $l$ ). The median estimated indices, estimated standard error of  $\hat{\lambda}_a^*$  ( $SE(\hat{\lambda}_a^*)$ ), the percentage relative standard error (RSE%) and mean square error (MSE) are also given.

$l$	$\hat{\lambda}_a$	$\hat{\lambda}_a^*$	(median) $\hat{\lambda}_a^*$	$SE(\hat{\lambda}_a^*)$	RSE%	MSE	CI-95 ( $\hat{\lambda}_a^*$ )
<b>age 1</b>							
1 cm	0.863	0.863	0.863	0.00910	0.000	0.0000	(0.86, 0.86)
2 cm	0.863	0.865	0.867	0.00939	1.085	0.00009	(0.84, 0.88)
3 cm	0.863	0.856	0.861	0.02803	3.274	0.00083	(0.80, 0.90)
4 cm	0.863	0.857	0.859	0.02791	3.257	0.00082	(0.81, 0.91)
5 cm	0.863	0.845	0.847	0.03694	4.370	0.00044	(0.81, 0.92)
6 cm	0.863	0.860	0.861	0.03514	4.088	0.00125	(0.79, 0.93)
7 cm	0.863	0.854	0.853	0.03803	4.454	0.00153	(0.80, 0.93)
<b>age 2</b>							
1 cm	6.496	6.496	6.491	0.00552	0.085	0.00003	(6.49, 6.50)
2 cm	6.496	6.486	6.486	0.02073	0.320	0.00053	(6.46, 6.53)
3 cm	6.496	6.504	6.506	0.05414	0.832	0.00299	(6.38, 6.60)
4 cm	6.496	6.498	6.500	0.05351	0.823	0.00287	(6.38, 6.60)
5 cm	6.496	6.514	6.517	0.07322	1.124	0.00567	(6.32, 6.65)
6 cm	6.496	6.503	6.507	0.07862	1.209	0.00623	(6.30, 6.65)
7 cm	6.496	6.486	6.491	0.09414	1.452	0.00897	(6.31, 6.64)
<b>age 3</b>							
1 cm	1.571	1.572	1.571	0.04499	2.861	0.00203	(1.49, 1.66)
2 cm	1.571	1.578	1.572	0.07268	4.605	0.00533	(1.45, 1.74)
3 cm	1.571	1.557	1.554	0.09893	6.353	0.00999	(1.41, 1.77)
4 cm	1.571	1.640	1.632	0.10687	6.517	0.00161	(1.38, 1.86)
5 cm	1.571	1.634	1.632	0.12411	7.593	0.01940	(1.31, 1.87)
6 cm	1.571	1.649	1.643	0.13337	8.086	0.02390	(1.31, 1.93)
7 cm	1.571	1.748	1.740	0.14741	8.432	0.05300	(1.28, 2.06)
<b>age 4</b>							
1 cm	1.584	1.581	1.581	0.06670	4.219	0.00446	(1.45, 1.71)
2 cm	1.584	1.597	1.596	0.11810	7.397	0.01410	(1.35, 1.83)
3 cm	1.584	1.613	1.619	0.14715	9.123	0.02250	(1.25, 1.89)
4 cm	1.584	1.563	1.568	0.14581	9.326	0.02170	(1.30, 1.84)
5 cm	1.584	1.586	1.581	0.15534	9.794	0.02410	(1.30, 1.90)
6 cm	1.584	1.596	1.595	0.16125	10.104	0.02620	(1.26, 1.93)
7 cm	1.584	1.502	1.500	0.16988	11.311	0.03550	(1.33, 1.83)
<b>age 5</b>							
1 cm	0.742	0.746	0.751	0.06817	9.1440	0.00466	(0.61, 0.87)
2 cm	0.742	0.738	0.729	0.10457	14.170	0.01100	(0.58, 0.96)
3 cm	0.742	0.765	0.756	0.11506	15.040	0.01380	(0.53, 1.00)
4 cm	0.742	0.764	0.757	0.11686	15.299	0.01410	(0.54, 1.00)
5 cm	0.742	0.801	0.787	0.12230	15.270	0.01840	(0.55, 1.07)
6 cm	0.742	0.779	0.765	0.11546	14.817	0.01470	(0.58, 1.02)
7 cm	0.742	0.828	0.814	0.12779	15.435	0.02360	(0.54, 1.11)
<b>age 6+</b>							
1 cm	1.074	1.073	1.065	0.06707	6.251	0.00450	(0.95, 1.20)
2 cm	1.074	1.067	1.060	0.10236	9.595	0.01050	(0.90, 1.28)
3 cm	1.074	1.036	1.028	0.11648	11.247	0.01510	(0.90, 1.26)
4 cm	1.074	1.009	1.003	0.11837	11.735	0.01830	(0.90, 1.25)
5 cm	1.074	0.950	0.944	0.11578	12.184	0.02880	(0.96, 1.19)
6 cm	1.074	0.944	0.930	0.11745	12.446	0.03090	(0.95, 1.20)
7 cm	1.074	0.913	0.905	0.11462	12.553	0.03910	(1.00, 1.14)

Table S7.3: Estimated abundance ( $\hat{\lambda}_a$ ) for saithe from the original data in year 2017 Q3 compared with estimated abundance ( $\hat{\lambda}_a^*$ ) from the reduced data for different sampling procedures of length groups ( $l$ ).

$l$	$\hat{\lambda}_a$	$\hat{\lambda}_a^*$	(median) $\hat{\lambda}_a^*$	SE( $\hat{\lambda}_a^*$ )	RSE%	MSE	CI-95 ( $\hat{\lambda}_a^*$ )
<b>age 0</b>							
1 cm	0.282	0.282	0.282	0.00000	0.00	0.00000	(0.28, 0.28)
2 cm	0.282	0.282	0.282	0.00000	0.00	0.00000	(0.28, 0.28)
3 cm	0.282	0.289	0.295	0.00626	2.17	0.00008	(0.28, 0.29)
4 cm	0.282	0.290	0.295	0.00592	2.04	0.00010	(0.28, 0.29)
5 cm	0.282	0.282	0.282	0.00000	0.00	0.00000	(0.28, 0.28)
6 cm	0.282	0.297	0.295	0.01022	3.44	0.00030	(0.28, 0.31)
7 cm	0.282	0.290	0.295	0.00594	2.05	0.00010	(0.28, 0.29)
<b>age 1</b>							
1 cm	0.123	0.123	0.123	0.00000	0.00	0.00000	(0.12, 0.12)
2 cm	0.123	0.123	0.123	0.00000	0.00	0.00000	(0.12, 0.12)
3 cm	0.123	0.117	0.111	0.00626	5.36	0.00008	(0.11, 0.12)
4 cm	0.123	0.118	0.115	0.00673	5.71	0.00008	(0.11, 0.13)
5 cm	0.123	0.125	0.123	0.00139	1.12	0.000003	(0.12, 0.13)
6 cm	0.123	0.112	0.114	0.01059	9.46	0.00024	(0.11, 0.13)
7 cm	0.123	0.116	0.114	0.00628	5.43	0.00009	(0.11, 0.13)
<b>age 2</b>							
1 cm	0.929	0.930	0.923	0.16851	18.13	0.02840	(0.64, 1.28)
2 cm	0.929	0.916	0.861	0.28468	31.06	0.08120	(0.55, 1.53)
3 cm	0.929	0.966	0.902	0.33158	34.32	0.11000	(0.53, 1.71)
4 cm	0.929	0.955	0.900	0.31885	33.38	0.10200	(0.49, 1.66)
5 cm	0.929	0.992	0.942	0.31609	31.85	0.10400	(0.48, 1.75)
6 cm	0.929	0.966	0.893	0.36374	37.66	0.13400	(0.47, 1.83)
7 cm	0.929	0.989	0.933	0.34996	35.40	0.12600	(0.45, 1.80)
<b>age 3</b>							
1 cm	11.238	11.270	11.249	0.53506	4.75	0.28700	(10.19, 12.30)
2 cm	11.238	11.179	11.187	0.92312	8.26	0.85600	(9.57, 13.11)
3 cm	11.238	11.109	11.082	1.07691	9.69	1.18000	(9.30, 13.27)
4 cm	11.238	11.000	11.009	1.08989	9.91	1.24000	(9.21, 13.15)
5 cm	11.238	10.891	10.871	1.11346	10.22	1.36000	(9.41, 13.03)
6 cm	11.238	10.920	10.905	1.06856	9.79	1.24000	(9.46, 13.04)
7 cm	11.238	10.840	10.839	1.08304	9.99	1.33000	(9.53, 13.05)
<b>age 4</b>							
1 cm	12.789	12.757	12.754	0.52780	4.14	0.28000	(11.79, 13.73)
2 cm	12.789	12.816	12.827	0.93741	7.31	0.87900	(10.76, 14.60)
3 cm	12.789	12.863	12.856	1.06438	8.27	1.14000	(10.68, 14.93)
4 cm	12.789	12.950	12.954	1.09842	8.48	1.23000	(10.56, 15.14)
5 cm	12.789	13.096	13.087	1.12912	8.62	1.37000	(10.51, 15.31)
6 cm	12.789	13.061	13.051	1.08819	8.33	1.26000	(10.42, 15.11)
7 cm	12.789	13.176	13.187	1.09385	8.30	1.35000	(10.33, 15.18)
<b>age 5</b>							
1 cm	2.971	2.971	2.966	0.13399	4.51	0.01800	(2.72, 3.24)
2 cm	2.971	3.048	3.037	0.27486	9.02	0.08150	(2.52, 3.62)
3 cm	2.971	3.000	2.974	0.31856	10.62	0.10200	(2.42, 3.65)
4 cm	2.971	3.038	3.005	0.34723	11.43	0.12500	(2.40, 3.77)
5 cm	2.971	2.971	2.968	0.33433	11.25	0.11200	(2.35, 3.64)
6 cm	2.971	2.980	2.964	0.38418	12.89	0.14800	(2.28, 3.77)
7 cm	2.971	2.940	2.922	0.39677	13.49	0.15800	(2.32, 3.76)
<b>age 6+</b>							
1 cm	2.819	2.818	2.820	0.05409	1.92	0.00293	(2.71, 2.92)
2 cm	2.819	2.787	2.784	0.08700	3.12	0.00860	(2.68, 2.96)
3 cm	2.819	2.808	2.808	0.11451	4.08	0.01320	(2.60, 3.04)
4 cm	2.819	2.800	2.795	0.12424	4.44	0.01580	(2.61, 3.06)
5 cm	2.819	2.793	2.791	0.13520	4.84	0.01890	(2.58, 3.07)
6 cm	2.819	2.814	2.823	0.14353	5.10	0.02060	(2.54, 3.10)
7 cm	2.819	2.800	2.794	0.16239	5.80	0.02670	(2.55, 3.14)