INTERBENCHMARK PROTOCOL ON ASSESSMENT MODEL CHANGES FOR COD (GADUS MORHUA) IN SUBAREAS 1 AND 2 (NORTHEAST ARCTIC) (IBPNEACOD2019)

VOLUME 1 | ISSUE 26

ICES SCIENTIFIC REPORTS

RAPPORTS SCIENTIFIQUES DU CIEM

ICES INTERNATIONAL COUNCIL FOR THE EXPLORATION OF THE SEA
CIEM CONSEIL INTERNATIONAL POUR L'EXPLORATION DE LA MER

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

H.C. Andersens Boulevard 44-46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

ISSN number: 2618-1371 I © 2019 International Council for the Exploration of the Sea

ICES Scientific Reports

Volume 1 | Issue 26

INTERBENCHMARK PROTOCOL ON ASSESSMENT MODEL CHANGES FOR COD (GADUS MORHUA) IN SUBAREAS 1 AND 2 (NORTHEAST ARCTIC) (IBPNEACOD2019)

Recommended format for purpose of citation:

ICES. 2019. Interbenchmark Protocol on assessment model changes for Cod (Gadus morhua) in subareas 1 and 2 (Northeast Arctic) (IBPNEACod2019).
ICES Scientific Reports. 1:26. 26 pp. http://doi.org/10.17895/ices.pub.5278

Editors

Daniel Howell

Authors

Hoskuldur Bjornsson • Bjarte Bogstad • Yury Kovalev • Paul Regular • Dimitri Vasilev • Natalia Yaragina

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

Contents

i Executive summary

IBPNEACOD 2019 was chaired by Daniel Howell from Norway and run in order to examine a proposed change to the North East Arctic (NEA) cod SAM model settings for variance in the data.

The SAM model was adopted at IBPCOD 2017, based on data up to 2016. Since that time the two large year classes from 2003 and 2004 have continued to age and continuing fishing with moderate F has led to good survivorship of these year classes. As a result, the age structure of the cod population has continued to improve. This means that several of the assumptions made at the most recent benchmark (IBPCOD 2017) about how to handle the older (age 12 and upwards) fish in the model may no longer be valid. One key assumption made was that the variance in the data was assumed constant for each tuning fleet. This assumption was examined at the 2019 Arctic Fisheries Working Group (AFWG) meeting and found to be unrealistic, with larger variances for many older age groups. Furthermore, these age groups now contain a significant biomass of fish. Consequently, a revised model setting was proposed, with a more detailed variance structure to account for this.

Two reviewers were appointed to examine the proposal. Both reviewers concurred that there was an issue that needed addressing, and that revising the variance structure was an appropriate response. However, both reviewers had concerns about the lack of justification for the specific choice of ages to select for different variances in the proposed approach. One reviewer was prepared to accept the changes with a further justification of the choice, while the other would only recommend an alternative approach (not fully investigated due to time constraints) with a simpler structure. The relevant advice drafting group (ADGANW 2019) concurred with this second viewpoint, and was further concerned that the structure chosen this year may represent overfitting and might need changing in future years.

The overall decision was therefore to keep with the status quo in the SAM model until the variance structure (and other measures such as increasing the age range of the tuning data) could be more thoroughly addressed at a full benchmark. IBPNEACOD 2019 strongly recommends that such a benchmark be held as soon as practical – likely early in 2021.

ii Expert group information

Expert group name	Interbenchmark Protocol on assessment model changes for Cod (*Gadus morhua*) in subareas 1 and 2 (Northeast Arctic) (IBPNEACod2019)
Expert group cycle	NA
Year cycle started	2019
Reporting year in cycle	1/1
Chair	Daniel Howell, Norway
Meeting venue and dates	By correspondence during May 2019

iii Term of reference

Term of reference	Addressed in this report
a) Review deviations from the stock annex assessment model setup proposed by the Arctic Fisheries Working Group (AFWG) 2019.	Yes

Stocks	Stock leader	Stock assessor
Cod (*Gadus morhua*) in subareas 1 and 2 (Northeast Arctic)	Yuri A. Kovalev	Yuri A. Kovalev and Bjarte Bogstad

1 Working Document 1. Proposed changes in the northeast Arctic cod stock assessment, 2019.

1.1 Background

Although North-East Arctic has never been driven to extremely low levels, it did experience a multidecadal period of high fishing pressures and moderately low stock biomass. As a result, the age structure became truncated, with few fish above age ca. 10 in the stock. This cod stock then had an extended period of management with moderate fishing pressure with a series of formal HCRs since 2002. Fishing pressure was cut in half from 2006 to 2008 and has stayed low since then; this reduction came later than the HCR introduction mainly due to enforcement issues. This period of low fishing pressure coincided with two very good year classes in 2004 and 2005. As a result of the good survivorship of these year classes, the cod stock biomass has drastically increased, and now appears to be fluctuating with recruitment variation rather than fishing intensity. More recently the age structure has also begun to improve (greater range of ages, more old fish) as these two year classes began to age beyond that seen in quantity in the recent stock history. Because this age-structure recovery has been led by this pulse of recruitment, the changes in the age structure have been rather rapid. For example, the fraction of the catch that is from age 10+ or 12+ fish has increased rapidly over the past few years (Figures 1.1 and 1.2.). The increase in the total stock numbers and biomass which is 10+ or 12+ (Figure 1.3) as well as these age-groups' proportion of the total numbers/biomass is also strong (Figures 1.4 and 1.5). Figure 1.5 also shows the ratio of 12+ biomass in catch compared to in stock for both assessments.

While such changes are of course desirable, they place challenges on the modelling and assessment of the stock. In this case, the choices made at the last interbenchmark about the age range for which variances could be estimated were valid at the time, but have been invalidated by the rapid change in age structure of the stock. The last interbenchmark was held prior to the 2017 stock assessment and thus used 2015 as the last data year for catches. Thus this year's assessment had three more years of available data and the two strong year classes were age 13 and 14 in the last data year compared to age 10 and 11 in the dataset used at the interbenchmark.

The tuning data in the model goes to age 15 for the catch and age 12 for the surveys. One might expect that the data on the oldest fish in each dataset would be more uncertain than that for younger fish. However, there were very few fish in these older age categories, making estimating uncertainty difficult. Furthermore, the low quantity of fish of this age previously made the tuning data unnecessary for the assessment. However, these simplifications are no longer valid, these age classes are now numerous enough to matter and for separate estimates of variance to be made. This document presents a proposal to adjust the SAM assessment model to account for these changes. We note that because the changes are ongoing, this issue will need revisiting in a more considered fashion once the large year classes have worked their way further through the population. We are therefore proposing the changes described here as an interim solution prior to a full benchmark, ideally just prior to the 2021 stock assessment. We would also highlight that the changes here do not affect the validity of the HCR or the biomass reference point. The changes are only significant in the more recent years, which correspond to the period of high biomass. Estimates of B_{lim}/B_{pa} are thus unaffected by these changes. Although estimates of F_{MSY} would likely be slightly affected, this quantity is not used in the HCR used for the management of this stock.

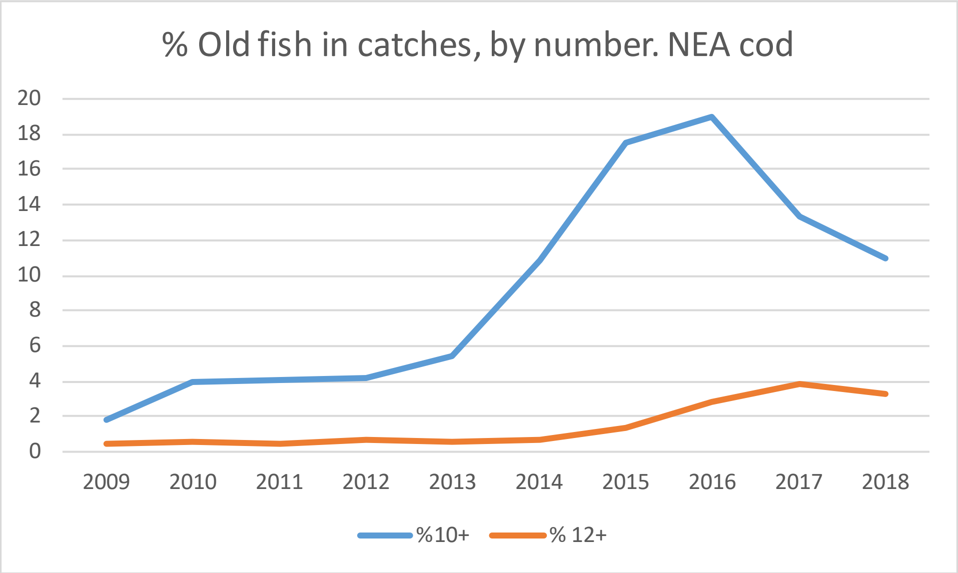


Figure 1.1. % old (ages 10+ and ages 12+) fish in catches from the fishery, by number.

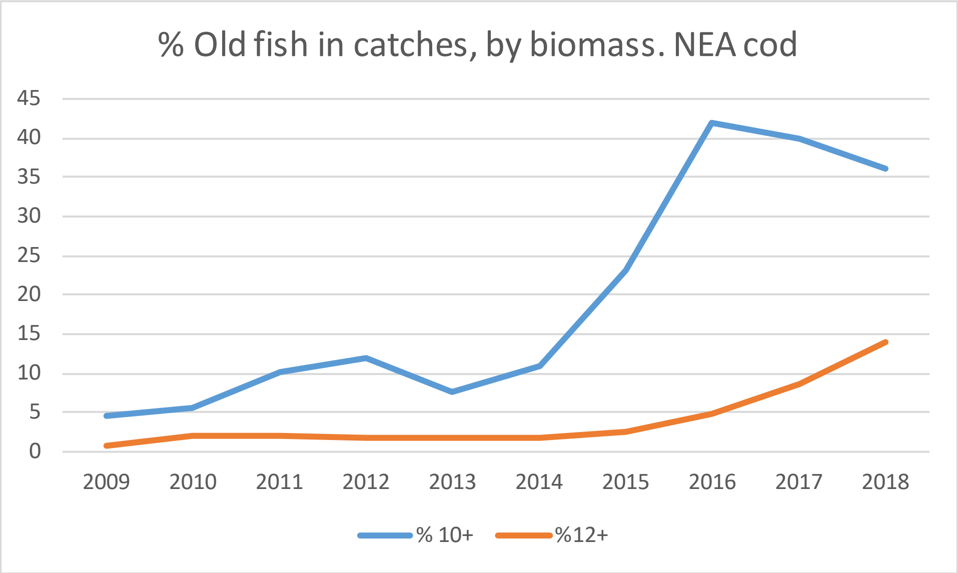


Figure 2.2. % old (ages 10+ and ages 12+) fish in catches, in biomass

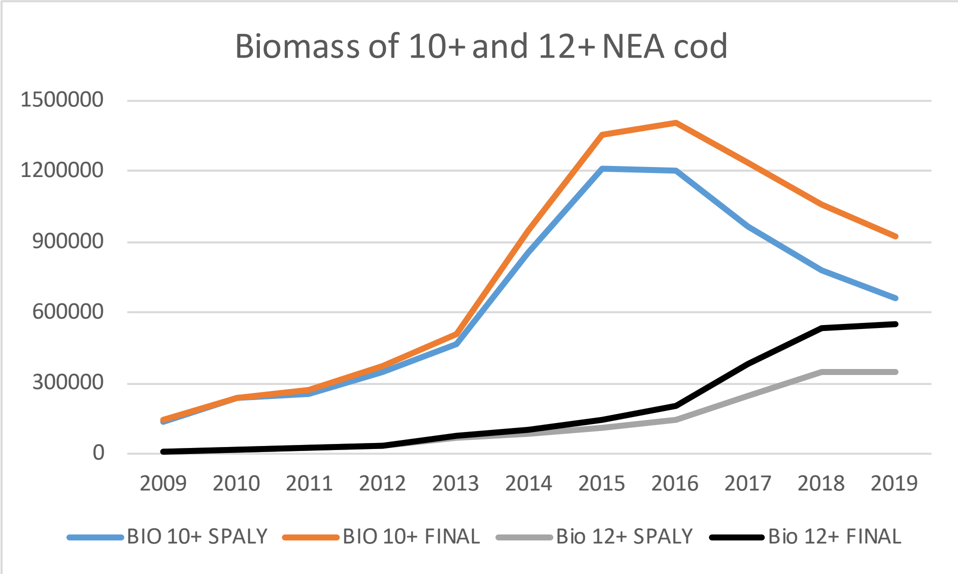


Figure 1.3. Age 10+ and 12+ biomass (tonnes) with FINAL and SPALY run

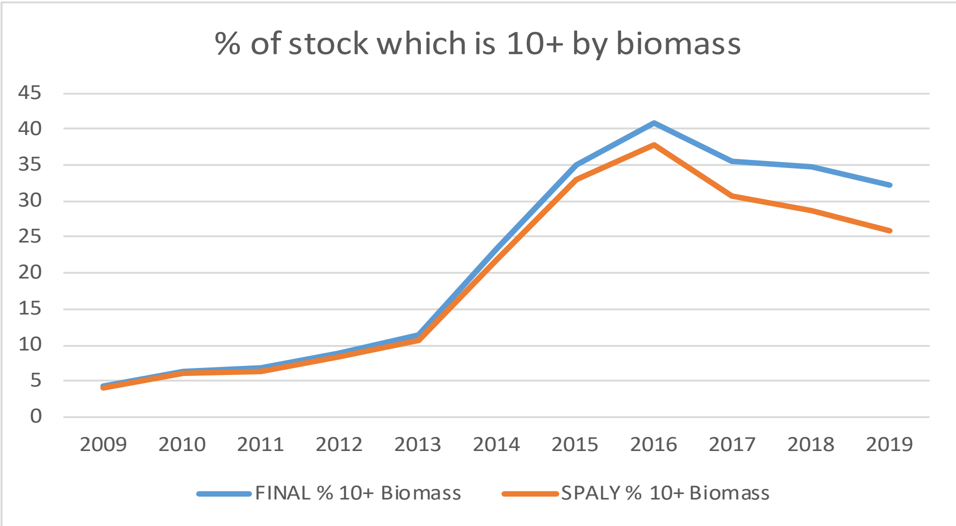


Figure 1.4. % age 10+ fish in stock, by biomass

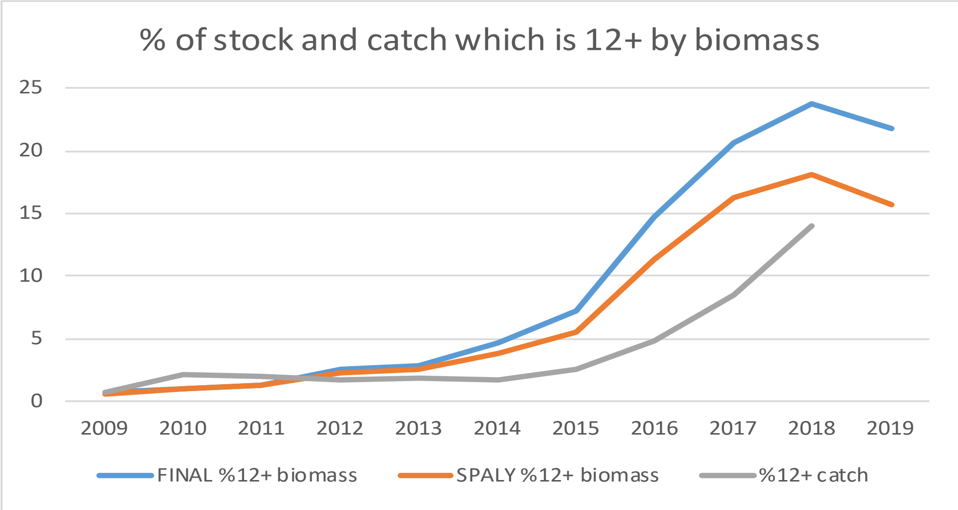


Figure 1.5. % age 12+ fish in stock, by biomass and in catch

1.2 Explanation to changes in SAM model settings for cod

In northeast Arctic Cod stock assessment using SAM the fleet-specific variance parameters for the observations are used, but within the same fleet these parameters are age-independent (see Table 1.1, representing the corresponding block in the model configuration file used in 2017 and 2018 following the IBP in 2017, hereafter denoted as SPALY/SPALY run).

The first row in the table corresponds to catch-at-age, further rows correspond to "fleet 15" (Barents Sea bottom trawl winter) "fleet 16" (Barents Sea acoustic winter+Lofoten), "fleet 18" (Russian late autumn survey) and "fleet 007" (ecosystem-survey). Note that for fleet 18 and 007, no data for the last year were available. The meaning of the same values in each of the rows is that the same values of variance are used for all age groups for a given fleet.

Table 1.

Coupling of the variance parameters for the observations - SPALY

0	0	0	0	0	0	0	0	0	0	0	0	0
-1	1	1	1	1	1	1	1	1	1	-1	-1	-1
-1	2	2	2	2	2	2	2	2	2	-1	-1	-1
3	3	3	3	3	3	3	3	3	3	-1	-1	-1
4	4	4	4	4	4	4	4	4	4	-1	-1	-1

At the interbenchmark it was decided to have the same observation variance at age within each of the 4 tuning series and in the commercial catches, i.e. the simplest possible structure. At the time this seemed reasonable since there were very few fish in the oldest age groups (and note that time at the interbenchmark for checking settings was limited). Since the IBP however, the proportion of the stock and catch in the oldest age groups has increased and can no longer be treated as insignificant (ca. 15% of the catch in biomass in 2018 was from age 12+ fish compared to 3% in 2015, Figure 1.5). At the same time the model diagnostics were indicating that the variance in the data for these oldest fish was (unsurprisingly) different from the younger fish (Figure 1.6).

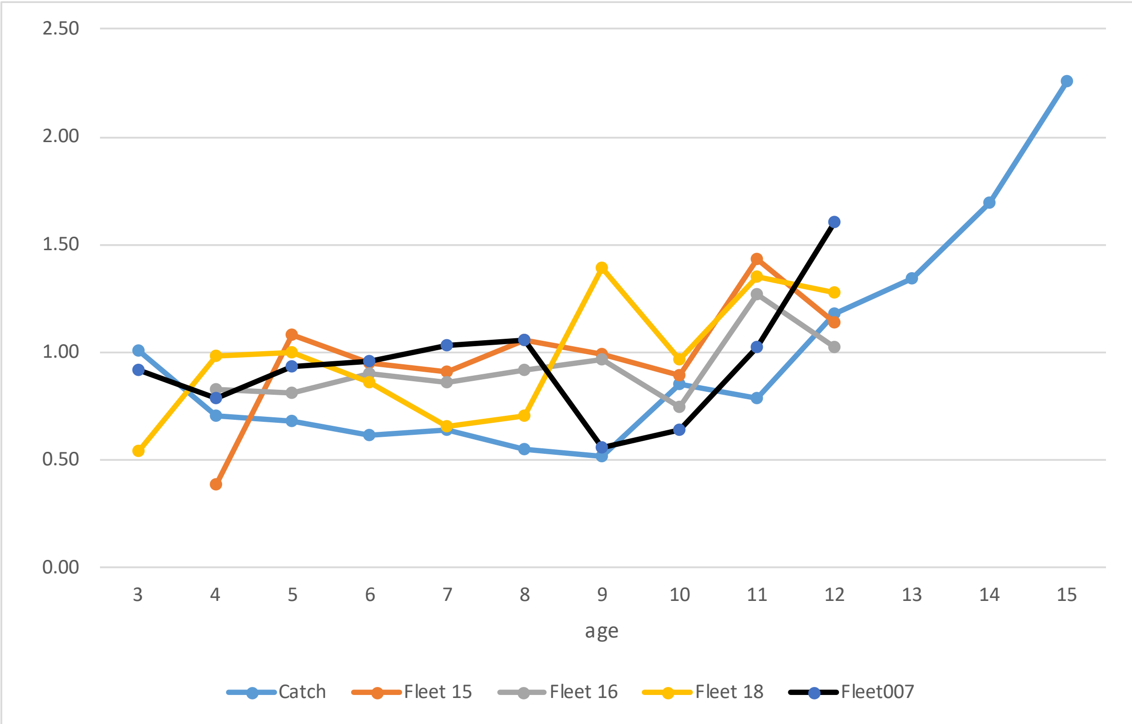


Figure 1.6. Residuals at age for the different data sources used in the assessment.

In WD_02 to AFWG-2019 it was proposed to use more parameters for coupling variances using as a starting point the values of standard deviations (STD) for the model residuals. It was proposed to use separate variances for ages with untypically high or small STD of residuals (Table 1.5) assuming that for such ages the use of fleet-average variance may not allow the model to describe the data adequately. Simultaneously it was intended not to include too many new free parameters into the model. The results presented in WD_02 showed better results of diagnostics (lower AIC value and better retrospective stability in terms of RHO(SSB,R,F)) with respect to the SPALY settings used at AFWG-2018.

The proposed in the WD setting of variance parameters for the observations are given in Table 1.2.

Table 1.2

Coupling of the variance parameters for the observations – WD02

0	0	0	0	0	0	0	0	0	0	0	5	5
-1	6	1	1	1	1	1	1	7	1	-1	-1	-1
-1	2	2	2	2	2	2	2	8	2	-1	-1	-1
3	3	3	3	3	3	9	3	9	9	-1	-1	-1
4	4	4	4	4	4	4	4	4	10	-1	-1	-1

At the meeting of AFWG-2019 the proposed changes to SAM settings were tested with the new dataset. Also a number of different variants of variance parameters settings, including enhanced and reduced number of age groups with coupled variance parameters, were tested.

The final choice was based on analysis of AIC values and measures of retrospective stability.

Table 1.3. Some criteria for diagnostic of fitting quality (Akaike criteria, Rho for F, R and SSB) for SAM runs with different assumption regarding ages with different variance of observations

	SPALY	WD-02	age 13 for catch added
AIC	2845	2568	2509
Rho(ssb)	8	10	8
Rho R	-27	-19	-20
Rho (F)	-3	-1	-1

According to Table 1.3, the case where age 13 (see table 4) was added to the block with ages 14 and 15 for catch data was found to be preferable as the estimated SD for residuals in the chosen blocks became closer to 1.0 (Tables 1.5 and 1.6) and model diagnostics (AIC values and retrospective measures) were improved. This adjusted model variance parameterization adds 6 model parameters, but improves the AIC score by 336.

The final adjusted variance structure, which is almost identical with the one proposed in WD02-2019, thus prescribes two variance parameters for catches (ages >12 and age 12 and above), three variance parameters for fleet 15 (separate for age four, separate for age 11, and common for the remaining ages), two variance parameters for fleet 16 (separate for age 11, common for remaining ages), two variance parameters for fleet 18 (one for ages 9, 11, and 12, and one common for the remaining ages), finally two parameters for fleet 07 (one for age 12 and one common for the remaining ages).

Table 1.4

Coupling of the variance parameters for the observations – FINAL run

0	0	0	0	0	0	0	0	0	0	5	5	5
-1	6	1	1	1	1	1	1	7	1	-1	-1	-1
-1	2	2	2	2	2	2	2	8	2	-1	-1	-1
3	3	3	3	3	3	9	3	9	9	-1	-1	-1
4	4	4	4	4	4	4	4	4	10	-1	-1	-1

Table 1.5. Standard deviations of residuals by fleets and age groups from SPALY run, in blue - SD for ages which later were dealt with as separate parameters. "Fleet 15" (Barents Sea bottom trawl winter) "fleet 16" (Barents Sea acoustic winter+Lofoten), "fleet 18" (Russian late autumn survey) and "fleet 007" (ecosystem-survey)

Age / Fleet	3	4	5	6	7	8	9	10	11	12	13	14	15
Catch	1.00	0.71	0.68	0.62	0.64	0.55	0.52	0.85	0.79	1.18	1.34	1.69	2.26
Fleet 15		0.38	1.08	0.95	0.91	1.06	0.99	0.90	1.44	1.14			
Fleet 16		0.83	0.81	0.90	0.86	0.92	0.97	0.75	1.27	1.03			
Fleet 18	0.54	0.98	1.00	0.86	0.66	0.70	1.39	0.97	1.35	1.27			
Fleet007	0.92	0.79	0.94	0.96	1.04	1.06	0.56	0.64	1.02	1.61			

Table 1.6. Standard deviations of residuals by fleets and age groups from FINAL run, dealing with SD for ages marked in blue as separate parameters. "Fleet 15" (Barents Sea bottom trawl winter) "fleet 16" (Barents Sea acoustic winter+Lofoten), "fleet 18" (Russian late autumn survey) and "fleet 007" (ecosystem-survey)

Age / Fleet	3	4	5	6	7	8	9	10	11	12	13	14	15
Catch	1.02	0.83	0.79	0.72	0.87	0.76	0.72	1.01	1.04	1.51	0.95	1.10	1.46
Fleet 15		0.64	0.84	0.99	0.89	1.06	1.14	1.05	0.95	1.25			
Fleet 16		0.89	0.95	0.80	0.89	0.97	1.06	0.86	1.06	1.04			
Fleet 18	0.76	1.16	1.02	0.82	0.80	0.84	1.01	1.49	1.05	1.11			
Fleet007	1.02	0.83	0.95	1.00	1.04	0.99	0.66	0.76	1.29	1.04			

The WG propose to use the above mentioned new SAM settings for NEA cod stock assessment as it gives apparently better results of retrospective diagnostics and better model fit to data (in terms of AIC) with respect to the SPALY SAM settings.

1.3 Differences in diagnostics and results

Figures 1.7 and 1.8 show the diagnostics for the two runs and Figure 1.9 shows the retro pattern

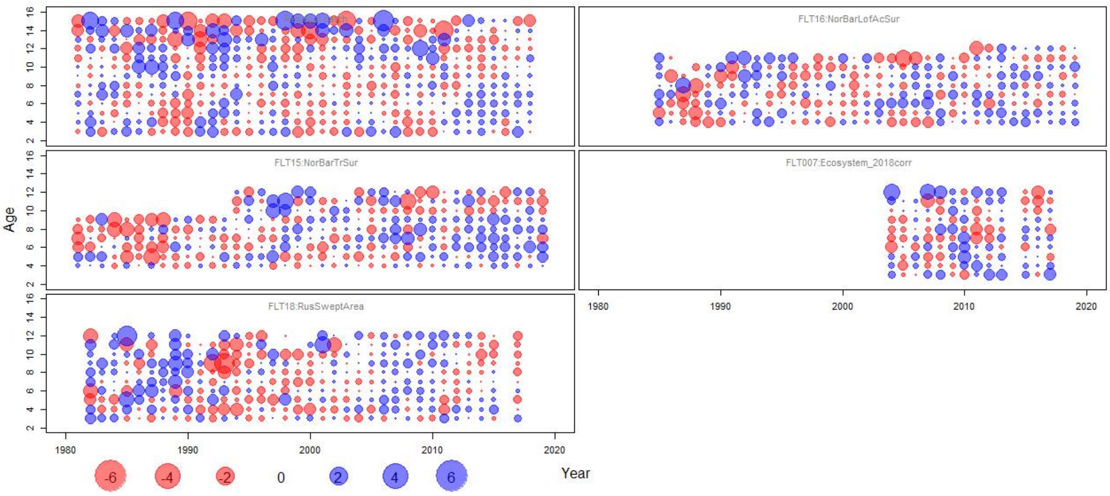


Figure 1.7. Log-catchability residuals of catch and fleets. SPALY run

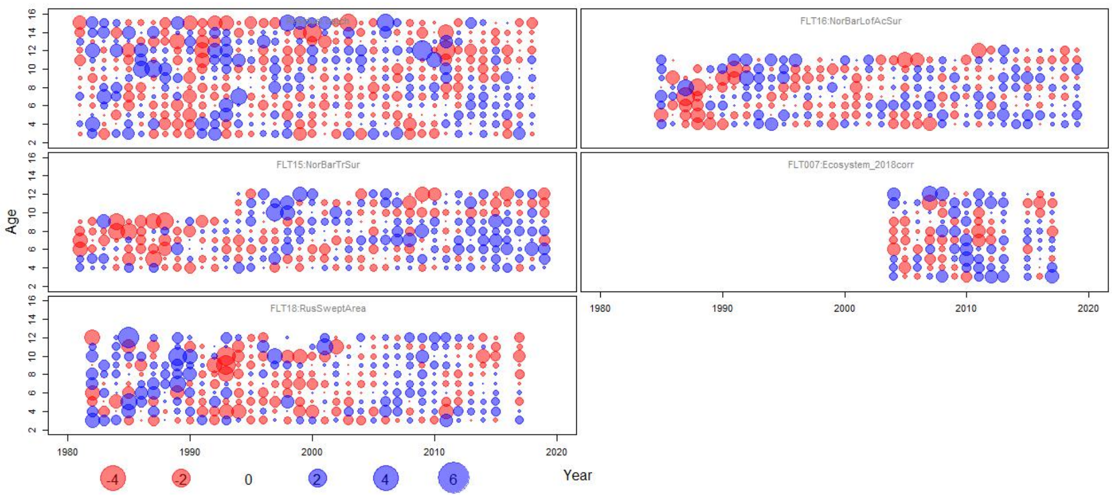
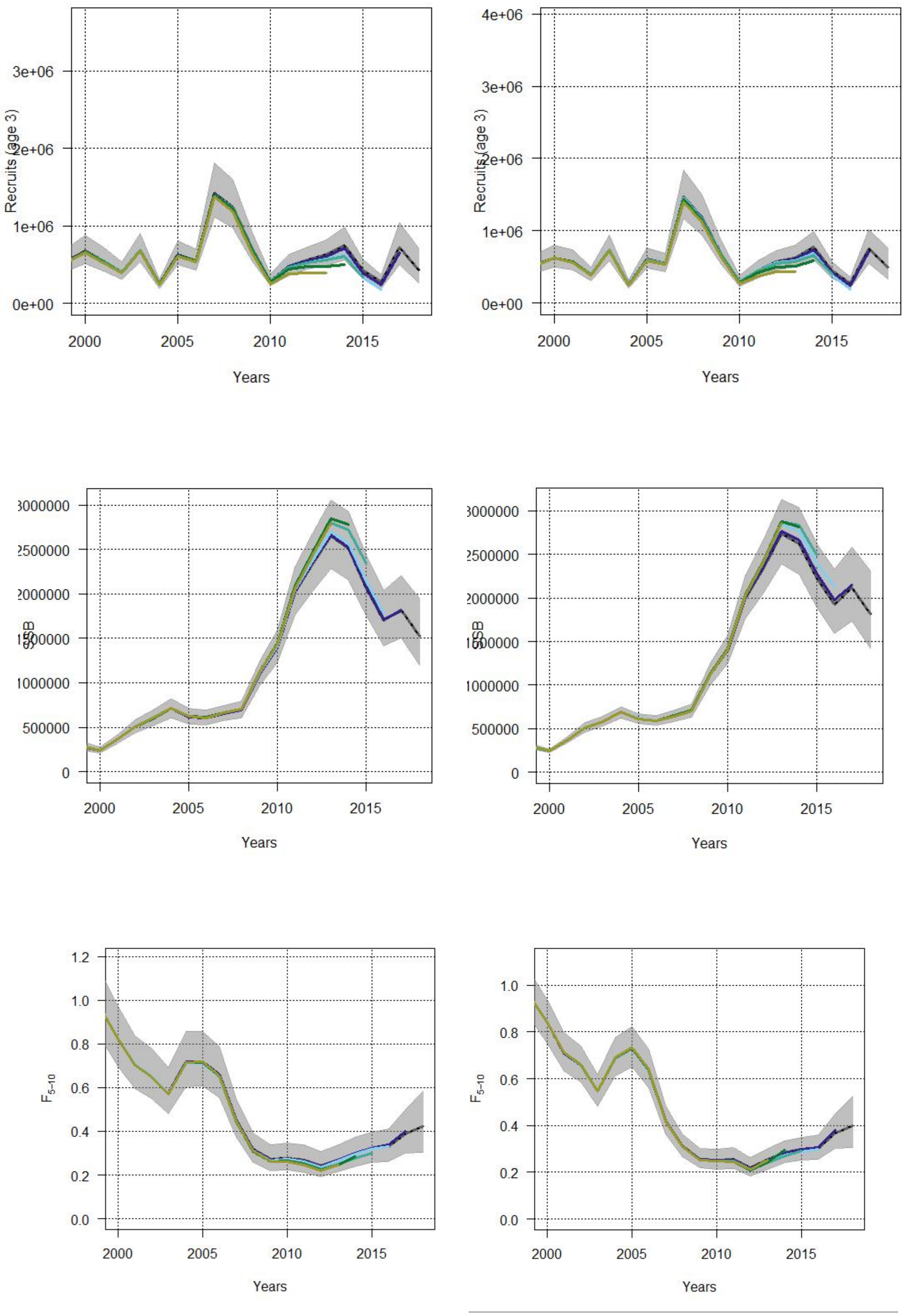


Figure 1.8. Log-catchability residuals of catch and fleets. Final run.


Recruits (age 3)
0e+00
1e+06
2e+06
3e+06
4e+06
SSB
0
500000
1000000
1500000
2000000
2500000
3000000
F_{5-10}
0.0
0.2
0.4
0.6
0.8
1.0
1.2
2000
2005
2010
2015
Years

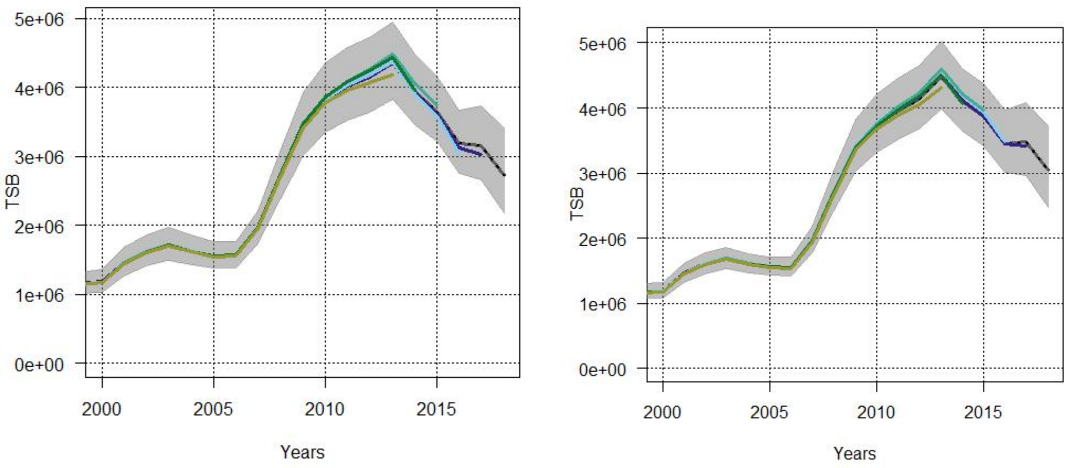


Figure. 1.9. Results of retrospective run for SPALY SAM settings (left) and Final SAM run (right)

1.4 Differences in results

The effect of new SAM settings on stock dynamic estimates is demonstrated in Figures 1.10 and 1.11.

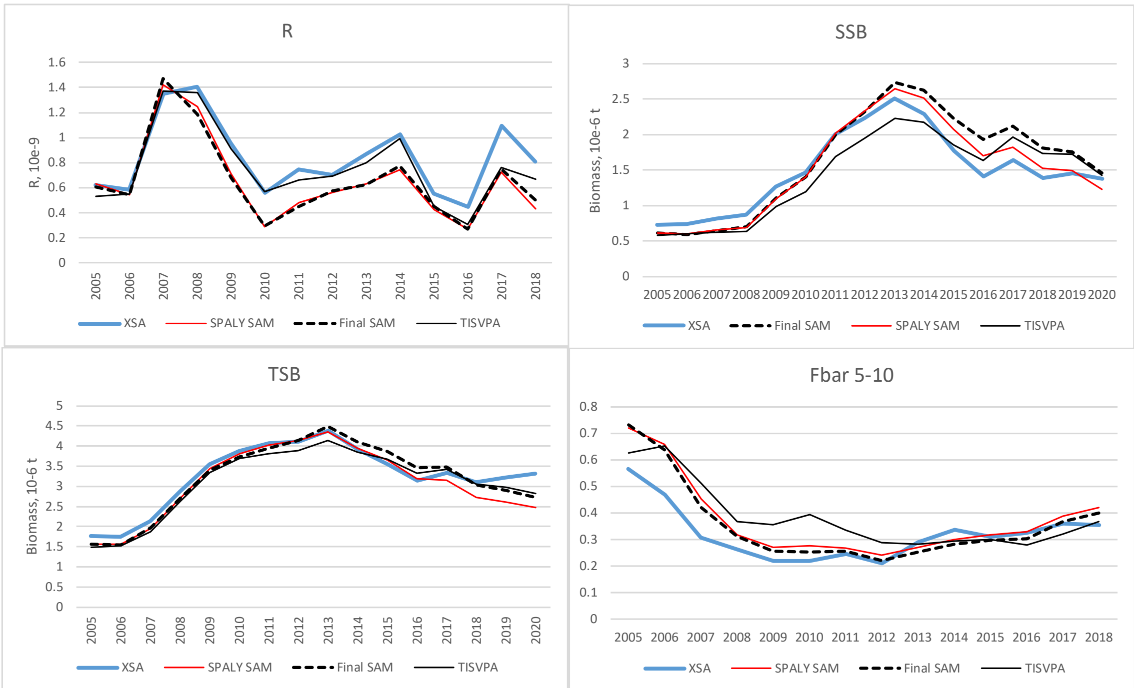


Figure. 1.10. Comparisons of R, SSB, TSB and F_{bar} dynamics from 4 models (SPALY SAM, Final SAM, XSA and TISVPA)

Figure. 1.11. Comparisons of numbers-at-age for 2019 from 4 models (SPALY SAM, Final SAM, XSA and TISVPA). Upper panel: Ages 3-15+, lower panel: Ages 9-15+

1.5 Concerns with the proposed assessment

Since the choice of which age groups should be linked will be strongly affected by the changing age structure in a recovering stock, the group believes that this needs evaluating and updating more frequently than at the benchmarks. AFWG (in common with many other groups) used to do something similar to the stock-size dependent catchability in XSA, and for much the same reason. There was considerable debate at the group with some concern raised that this increased the modelled stock size for ages without survey data. However, the final consensus was that the forced linked variances was clearly wrong, and that we should move away from this. This will be examined in more detail at the next benchmark.

The change in assessment settings resulted in an increase in the assessed stock biomass, especially for the oldest ages (13+) for which no tuning data were available. This change is being proposed for the 2019 assessment, with the rationale that this improved the model fitting (including on AIC), reduced most retros and misfits.

The estimated stock numbers for age groups 14 and 15+ in 2019 are the highest in the time-series and the 15+ value is twice the previous maximum for that age group. Of the SSB in 2019, 31% consists of age 13+. The model results show that the fishing mortality on these age groups in recent years is much lower than experienced in the late 1940s and early 1950s when catches of these age groups were high (i.e. higher than in recent years). The recent average selectivity for those age groups used in the prediction implies that the in the years to come, a considerable proportion of the 15+ group will be fish older than 15 years, a situation which has not been experienced before. Stock and fishery dynamics (mortality, selectivity, growth) for older cod, in particular age 15+, is largely unknown. There is a single commercial fleet with a dome shaped selectivity, and no survey data on the oldest fish. This leads to the possibility that the model could estimate an artificially low selectivity and artificially high stock for these oldest fish. Investigations at AFWG showed no evidence of this happening using the SPALY run (simulations showed the model seemed robust to this possibility, while examination of the age composition in fleet sector catching the oldest fish showed no major deviations between this validation dataset and the model results).

A simulation study indicates that the model does reliably estimate the fishing mortality-at-age over time in the current data situation. The estimation accuracy does naturally decrease for at ages where less data are available. This simulation did not study the potential effect of conflicting data sources.

It was noted that model estimates for age 15+ for the proposed final run may be going up faster than the 15+ catch data suggested, and the short available series of survey estimates for ages 13+ from FLT16 (not included in the tuning) also indicates less of an increase in age groups 13+ in 2017-2019 than the assessment indicates. (Figure 1.12.) FLT 16 includes coverage of the spawning area and is the one of the two surveys with data from the last year which is considered to cover the largest part of the 10/12+ stock. However, Figure 1.13 shows that the model results reflect the proportion of age 12+ fish found in the catches of the Norwegian conventional fleet. Thus, these two data sources show different trends in the last three years.

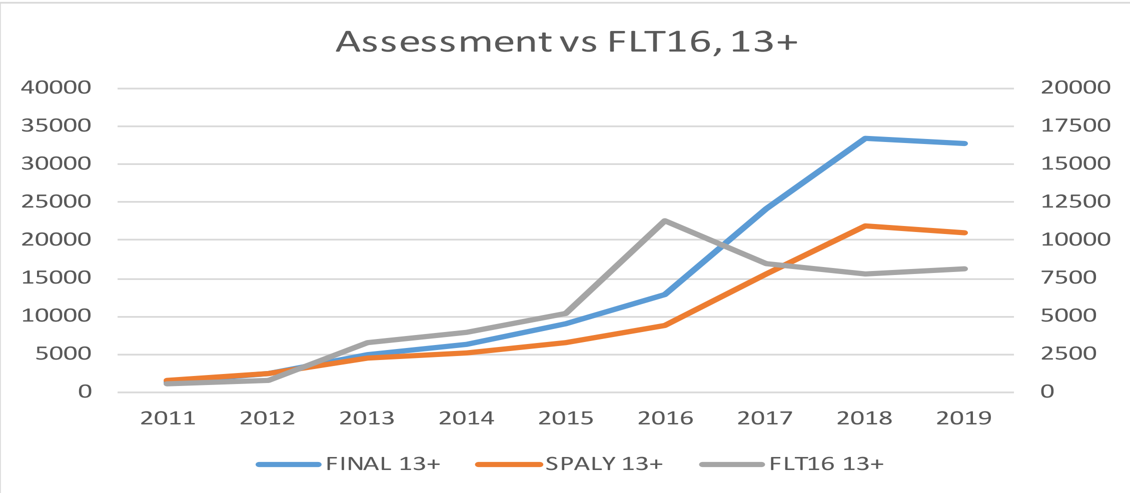


Figure 1.12. Number-at-age 13+ from FLT 16 (Barents Sea acoustic+Lofoten) vs. number-at-age from the two assessments.

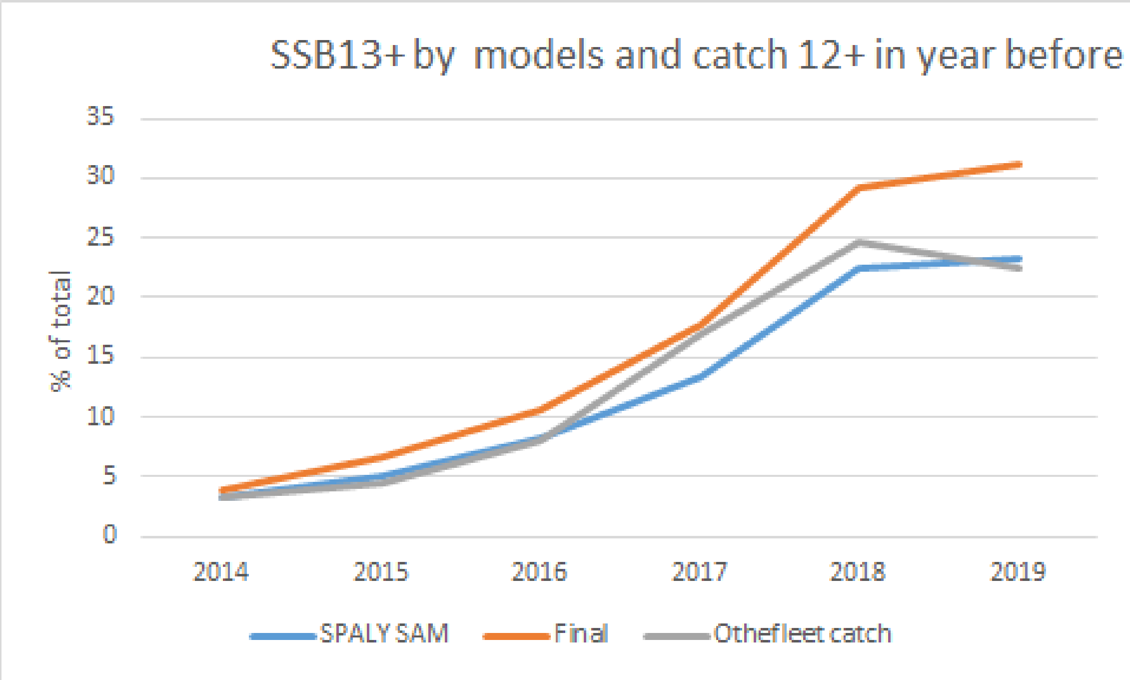


Figure 1.13. SSB of age 13+ vs. catch of age 12+ in the Norwegian conventional (non-trawl) fleet the previous year.

2 Review of the proposed changes to the northeast Arctic Cod stock assessment (Paul Regular)

16 May 2019

In their working paper, the authors outline some logical changes to the configuration of the SAM model used to assess the northeast Arctic cod stock. Given the trends in the data and patterns in the model diagnostics, changes to the model are justified in my humble opinion. Nevertheless, there are several aspects of the working paper that could benefit from more clarification, and more details on the merits of specific configurations could be provided.

My primary, and very minor, concern regarding the proposed changes lies with the very directed nature of the changes. Specifically, I am referring to the post hoc use of results from the SPALY run to modify the assumptions around age-specific observation error. I realize iterative modifications of assessment models like this are common, however, troublesome diagnostics ideally have an explanation and, with the potential explanation in mind, the model is adjusted accordingly. In the case of the proposed changes, it is reasonable to expect that observations of older fish may be more noisy than more abundant and more catchable younger age classes, however, I struggle to see how observations of 11 year old cod would be different from 12 year old cod. I wonder if the apparent differences, based on the post hoc analysis of the residuals, are an artefact of differing sample sizes? For instance, there are more than twice the number of years where age 11 fish were observed compared to age 12 fish for the Barents Sea acoustic winter + Lofoten survey (Fleet 16). To me it seems more realistic that the variance of these two age groups are similar and I would therefore like to see a run that couples the variance parameters for ages 11+ cod for each surveys. This would be a simpler modification that would be easier to explain and I suspect this change, like the FINAL run, will address the issues with the fits to the older ages.

On a related note, why are the estimates so tight for age 4 cod from the bottom trawl winter survey (Fleet 15)? Also, are there any ideas on what the issue is with observations of age 9 cod from the Russian survey (Fleet 18)? Is it driven by the large residuals from 1992 and 1993? If the latter is true, then perhaps the variance parameter for age 9 cod from the Russian survey could be coupled with ages 3-10.

In summary, I agree that changes are justified to address the residual patterns in older fish and I think that the approach used to change the configuration is logical but perhaps a little too specific. It would be great to see more detail on other model configurations (such as the simple suggestion above, if time permits) and some potential explanations of the patterns observed.

3 Northeast artic Cod review (Höskuldur Björnsson)

20 May 2019

North East Artic cod
Review

Höskuldur Björnsson

May 20th 2019

1 Introduction

Advice for Barentssea cod has since 2017 been based on a SAM assessment. The settings of the model are rather unconventional with 21 correlation parameters in the observation models estimated. table(1). The correlations structures estimated look reasonably sensible i.e more correlation between agegroups at higher age.

The correlation matrices shown here are for model 3 (NEAcod2019block)that is the alternative with the least AIC. Models 2 (NEAcod2019WD) and 3 have exactly the same number of observation variance parameters. One more parameter is used for the catches in model 3 but one less for FLT15 i.e the trawl survey in February). Generally the parameter setup for model 3 looks more sensible than for model 2, taking fleet 15 in model 2 variance parameter 3 is used for ages 5- 10 and then again for age 12 special parameter is set only for age 11. In model 3 the same observation variance is used for ages 4-10 for fleet 15 but model 2 has indicated that variance for age 4 was much lower than for the rest (0.4 vs 0.64). In model 3 the observation variance for fleet 15 is 0.56 for ages 4-10. (all these variances will have to be looked at taking correlations into account.)

Then we come to the main problem. We have plenty of fish of age 11 and older. The surveys end at age 12 but variance is estimated seperately for ages 10 and 11 for all fleets. And the estimated standard errors for ages 11 and 12 are high or 1.0085, 0.7190, 1.1780108, 0.6435177 for fleets 15, 16, 18 and 7. The lowest CV is from the ECOSYSTEM survey (fleet 7) that is not available at the time of assessment. For comparison CV on the landing for ages 11 and 12 is 0.23 and 0.6 for ages 13-15. The random walk constraint in F has standard error of

0.33 (same value for all ages). The estimated correlation parameter on changes in F is 1.357 meaning relatively much correlation between agegroups, reducing the effect of the random walk constraint. (1.357 is equivalent to 0.876 in a first order AR model).

Comparison of model 1 and model 3 2 shows the same general pattern of F i.e reducing after age 11. This reduction in F is questionable without some explanation of what might be causing it. Survey indices should be extended to "15+", at least for one of the surveys and the q constrained to be the same for some of the oldest age groups, at least for surveys considered to cover the whole stock. There should not be that much difference between age 11 and 15 in terms of catchability, the question is more what is causing less than predicted amount of old fish. The main problem is though that there is not sufficient information in the data to estimate development of the oldest agegroups.

An alternative run was attempted where the run called "NEAcod2019block" was changed so F on ages 11-15 was constrained to be the same. It changes the estimated spawning stock considerably, and the selection on the oldest fish is of course very different (figure 1). The negative log-likelihood changes from 1181.147 to 1183.934. The number of parameters is reduced by 3 so the AIC should be lower. The problem is that the F parameters (keyLogFsta in conf) are not included in the total number of parameters that are 70 in both cases instead of 82 and 79. The spawning stock in 2019 in run 3 is 35% higher than in the run with flat selection above age 10. The effect on adviced F are though less as the advice is based on F going down with age so even though the virtual fish contributes to the stock size it does not contribute to the advice. (Tested with the MUPPET program). This does though only apply if SSB_{2020} is above the trigger of $3 \times B_{pa} = 1380$ thous tonnes. It is not there in 2019 but might be in 2020.

One problem with the current assessment is exclusion of ages 1 and 2. They are of course problematic for example due to variability in M, part of which is included in cannibalism. The model used for assessment estimates some kind of variability in M, feature abused for the adult fish but could be useful in recruitment estimates. Estimates of recruitment are also needed in the projections as **the TAC is calculated as the average catch predicted for the coming 3 years, using the target level of exploitation (Ftr)**. Also **If the spawning-stock biomass in the present year, the previous year, and each of the three years of prediction is above Bpa, the TAC should not be changed by more than ±20% compared with the previous year's TAC. In this case, Ftr should however not be below 0.30.**. The advice is therefore based on projections but it does though look like Ftr is based on only the estimated SSB in the advice year.

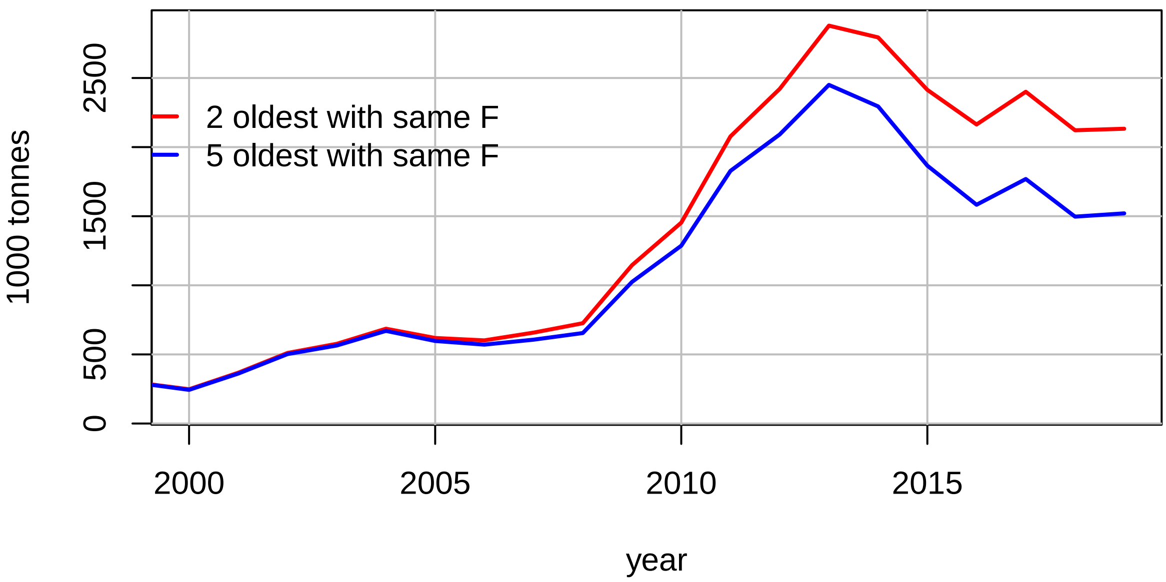


Figure 1: Comparison of SSB based on the run 3 and the same run with F constrained to be the same for 11-15. SSB before 2000 is identical.

	3	4	5	6	7	8	9	10	11	12	13	14
Residual catch												
FLT15:NorBarTrSur	-1	0	1	2	3	4	4	4	4	-1	-1	-1
FLT16:NorBarLofAcSur	-1	5	6	7	8	9	10	10	10	-1	-1	-1
FLT18:RusSweptArea	11	12	13	14	14	14	14	14	14	-1	-1	-1
FLT007:Ecosystem2018corr	15	16	17	18	19	20	20	20	20	-1	-1	-1

Table 1: Coupling of correlations in the SAM setup for NEA cod

	3	4	5	6	7	8	9	10	11	12	13	14
3	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
4	-1.00	1.00	0.30	0.07	0.02	0.01	0.01	0.01	0.01	-1.00	-1.00	-1.00
5	-1.00	0.30	1.00	0.25	0.07	0.04	0.03	0.03	0.03	-1.00	-1.00	-1.00
6	-1.00	0.07	0.25	1.00	0.29	0.16	0.14	0.12	0.10	-1.00	-1.00	-1.00
7	-1.00	0.02	0.07	0.29	1.00	0.55	0.48	0.42	0.37	-1.00	-1.00	-1.00
8	-1.00	0.01	0.04	0.16	0.55	1.00	0.87	0.76	0.67	-1.00	-1.00	-1.00
9	-1.00	0.01	0.03	0.14	0.48	0.87	1.00	0.87	0.76	-1.00	-1.00	-1.00
10	-1.00	0.01	0.03	0.12	0.42	0.76	0.87	1.00	0.87	-1.00	-1.00	-1.00
11	-1.00	0.01	0.03	0.10	0.37	0.67	0.76	0.87	1.00	-1.00	-1.00	-1.00
12	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
13	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
14	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00

Table 2: Estimated correlation matrix for FLT15:NorBarTrSur

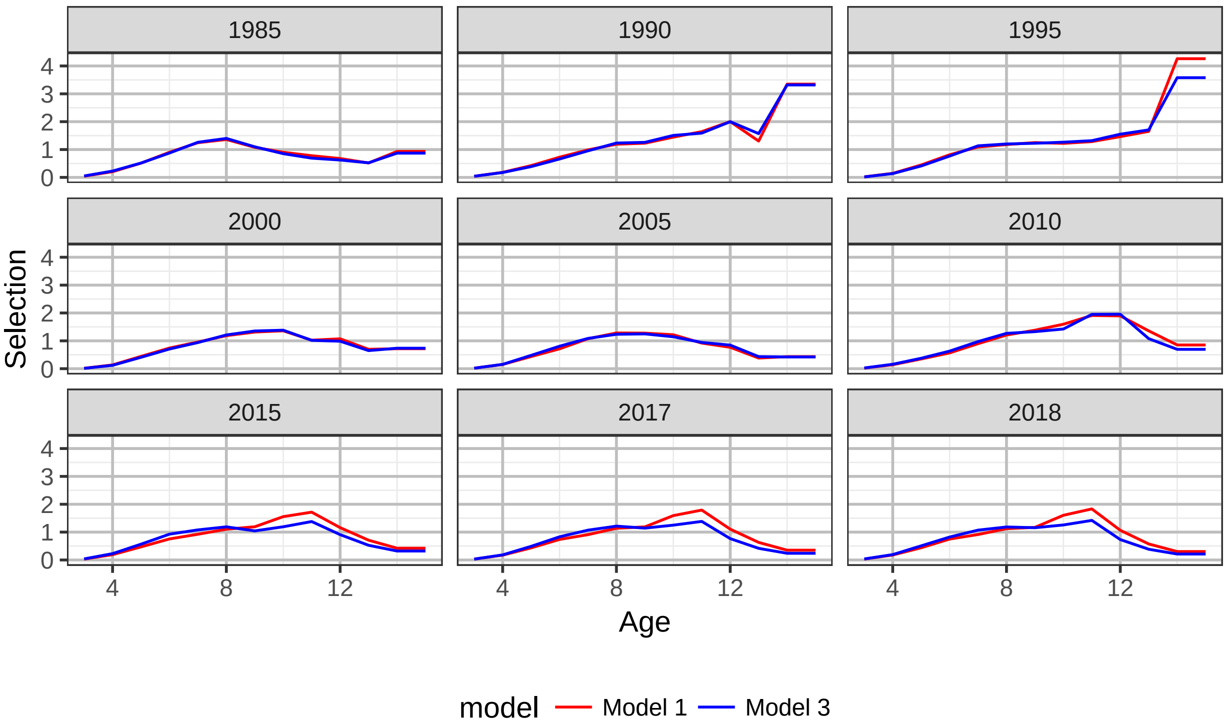


Figure 2: Selection by ageo with 5 years time interval

2 Alternative way to set variances

One way of estimating observation variance for the catches is to use pattern in residuals from Shephard - Nicholsson model $log(Cno)$ $factor(year) + factor(age) + factor(yearclass)$ and estimate one multiplier on it. Similarly the patten of residuals in the surveys could be obtained from an Adapt type model (model where CNO does take completely over historically) and one multiplier estimated for each survey.

The surveys give indications that number of fish in the oldest agegroups is increasing. The problem is that there were zero or very few fishes in those agegroups earlier, just a coincidence if 1, 2, 3 or 4 otholits were

	3	4	5	6	7	8	9	10	11	12	13	14
3	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
4	-1.00	1.00	0.31	0.07	0.04	0.03	0.03	0.03	0.03	-1.00	-1.00	-1.00
5	-1.00	0.31	1.00	0.21	0.12	0.11	0.10	0.10	0.10	-1.00	-1.00	-1.00
6	-1.00	0.07	0.21	1.00	0.57	0.52	0.45	0.47	0.48	-1.00	-1.00	-1.00
7	-1.00	0.04	0.12	0.57	1.00	0.91	0.80	0.82	0.84	-1.00	-1.00	-1.00
8	-1.00	0.03	0.11	0.52	0.91	1.00	0.88	0.90	0.92	-1.00	-1.00	-1.00
9	-1.00	0.03	0.10	0.45	0.80	0.88	1.00	0.98	0.96	-1.00	-1.00	-1.00
10	-1.00	0.03	0.10	0.47	0.82	0.90	0.98	1.00	0.98	-1.00	-1.00	-1.00
11	-1.00	0.03	0.10	0.48	0.84	0.92	0.96	0.98	1.00	-1.00	-1.00	-1.00
12	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
13	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
14	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00

Table 3: Estimated correlation matrix for FLT16:NorBarLofAcSur

	3	4	5	6	7	8	9	10	11	12	13	14
3	1.00	0.25	0.05	0.01	0.01	0.01	0.00	0.00	0.00	-1.00	-1.00	-1.00
4	0.25	1.00	0.19	0.06	0.04	0.03	0.02	0.01	0.01	-1.00	-1.00	-1.00
5	0.05	0.19	1.00	0.31	0.20	0.13	0.09	0.06	0.04	-1.00	-1.00	-1.00
6	0.01	0.06	0.31	1.00	0.65	0.43	0.28	0.18	0.12	-1.00	-1.00	-1.00
7	0.01	0.04	0.20	0.65	1.00	0.65	0.43	0.28	0.18	-1.00	-1.00	-1.00
8	0.01	0.03	0.13	0.43	0.65	1.00	0.65	0.43	0.28	-1.00	-1.00	-1.00
9	0.00	0.02	0.09	0.28	0.43	0.65	1.00	0.65	0.43	-1.00	-1.00	-1.00
10	0.00	0.01	0.06	0.18	0.28	0.43	0.65	1.00	0.65	-1.00	-1.00	-1.00
11	0.00	0.01	0.04	0.12	0.18	0.28	0.43	0.65	1.00	-1.00	-1.00	-1.00
12	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
13	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
14	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00

Table 4: Estimated correlation matrix for FLT18:RusSweptArea

obtained. A method that has been used in the Muppet model and was originally tested on the Norwegian Spring Spawning herring is to use modified logresiduals i.e $log\frac{(I+R)}{(\widehat{I}+R)}$ where R is a number correponding to 4 sampled otholits i.e when sampling error starts to dominate. What this method does is to increase the standard deviation where numbers in stock are low. By using this method surveys that are thought to cover older agegroups can be used without the estimated CV becoming too high. The estimated q in the surveys are shown in figure 3. What can be ween there are contratictory trends in the February surveys (F15 and F16) that need to be investigated.

The same feature can be seen in figure 4 but fleet 7 is becoming closer to flat and "flat q" might be obtained by weighted sum of F15 and F16.

	3	4	5	6	7	8	9	10	11	12	13	14
3	1.00	0.76	0.42	0.18	0.04	0.01	0.01	0.00	0.00	-1.00	-1.00	-1.00
4	0.76	1.00	0.55	0.24	0.05	0.02	0.01	0.01	0.00	-1.00	-1.00	-1.00
5	0.42	0.55	1.00	0.44	0.10	0.03	0.02	0.01	0.01	-1.00	-1.00	-1.00
6	0.18	0.24	0.44	1.00	0.22	0.07	0.04	0.03	0.02	-1.00	-1.00	-1.00
7	0.04	0.05	0.10	0.22	1.00	0.32	0.19	0.12	0.07	-1.00	-1.00	-1.00
8	0.01	0.02	0.03	0.07	0.32	1.00	0.60	0.37	0.22	-1.00	-1.00	-1.00
9	0.01	0.01	0.02	0.04	0.19	0.60	1.00	0.60	0.37	-1.00	-1.00	-1.00
10	0.00	0.01	0.01	0.03	0.12	0.37	0.60	1.00	0.60	-1.00	-1.00	-1.00
11	0.00	0.00	0.01	0.02	0.07	0.22	0.37	0.60	1.00	-1.00	-1.00	-1.00
12	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
13	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
14	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00

Table 5: Estimated correlation matrix for FLT007:Ecosystem2018corr

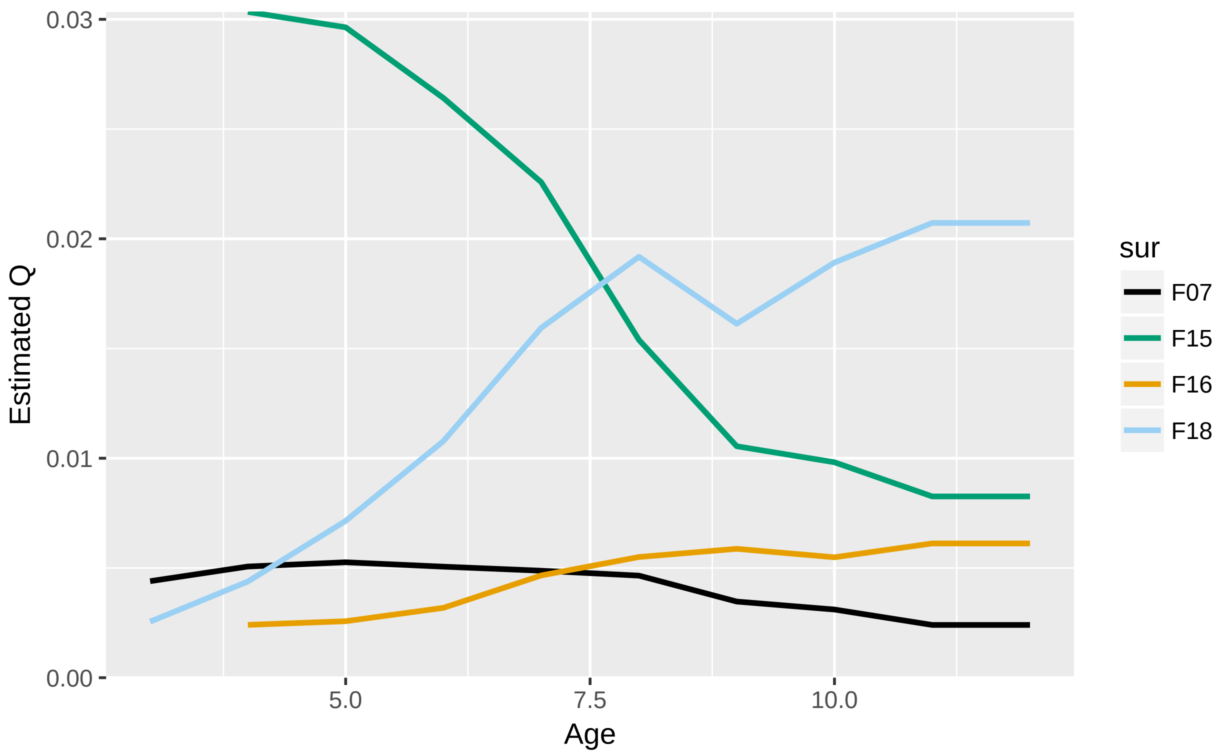


Figure 3: Estimated Q from the surveys based on Model 3

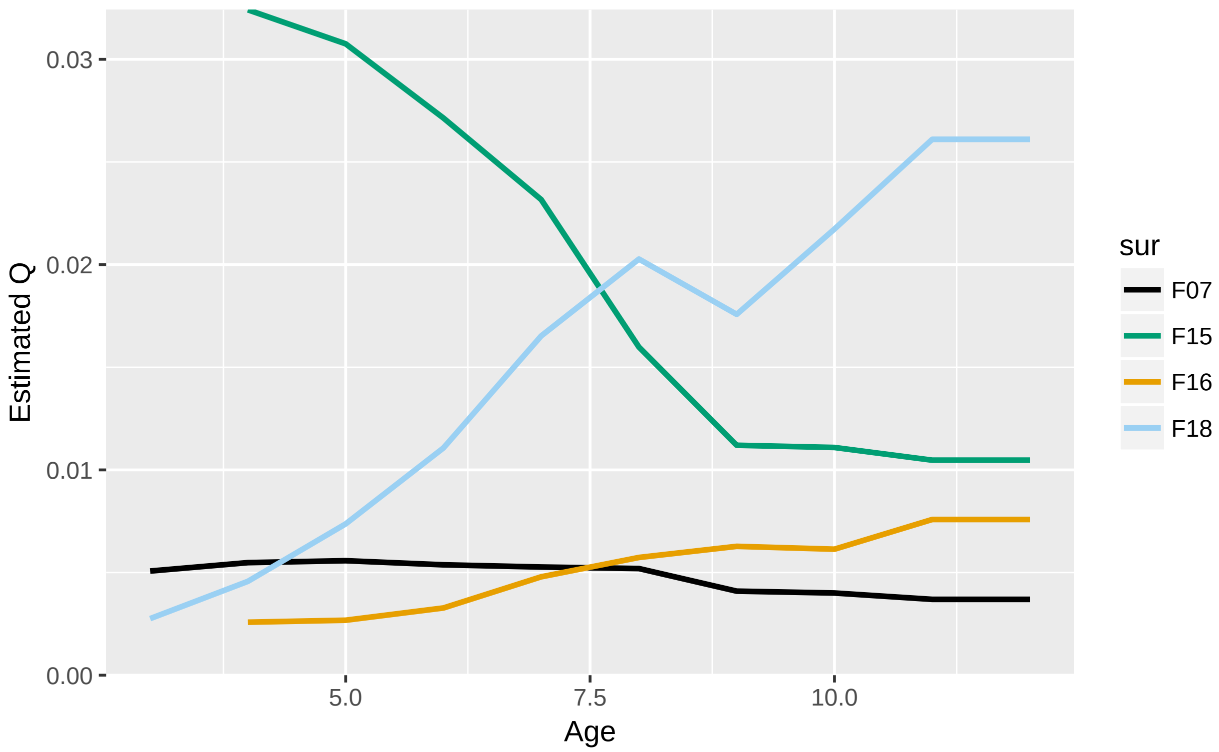


Figure 4: Estimated Q from the surveys based on Model 3 changed so F on last 5 agegroups is the same

3 Icelandic cod

The same problem shows up for Icelandic cod, i.e that fishing mortality by age is predicted to go down. As fishing mortality has been low for some years this is starting to have some effect on the predicted biomass and there is in addition some signal in the survey data for the older age groups. . When the number of agegroups usee in survey tuning is increased from 1-10 to 1-14 years the pattern of fishing mortality by age changes considerably (figure 5). The reference biomass changes by 7% and the difference in Q is most 9% in the March survey but 25% in the autumn survey. The F changes from boing down with age after age 9 to going slightly up.

TAC for Icelandic cod is a proportion of the reference biomass. For NEA cod estimated selection pattern is used to compile the catch so the effect on advice could potentially be less.

Another option here would be use flat selection but estimate q by age more freely. Also M used in the assessment might be looked at for the older fish. Neither having selection of the fisheries nor Q in one survey constrained to be flat for the older ages should not be an acceptable option.

Estimated observation variance for the surveys for Icelandic cod are much lower than for NEA cod. SAM estimates for Icelandic cod (variance the same for all age groups) are 0.25 for the March survey, 0.28 for the October survey. For the bottom trawl survey in February model 3 gives CV=0.56 for 4-10 bu 1.0 for 11-12. The only survey with reasonably low CV is the the ecosystem survey where estimated CV for ages 3-10 is 0.28. Treating F15 and F16 as separate indices might cause the problem here.

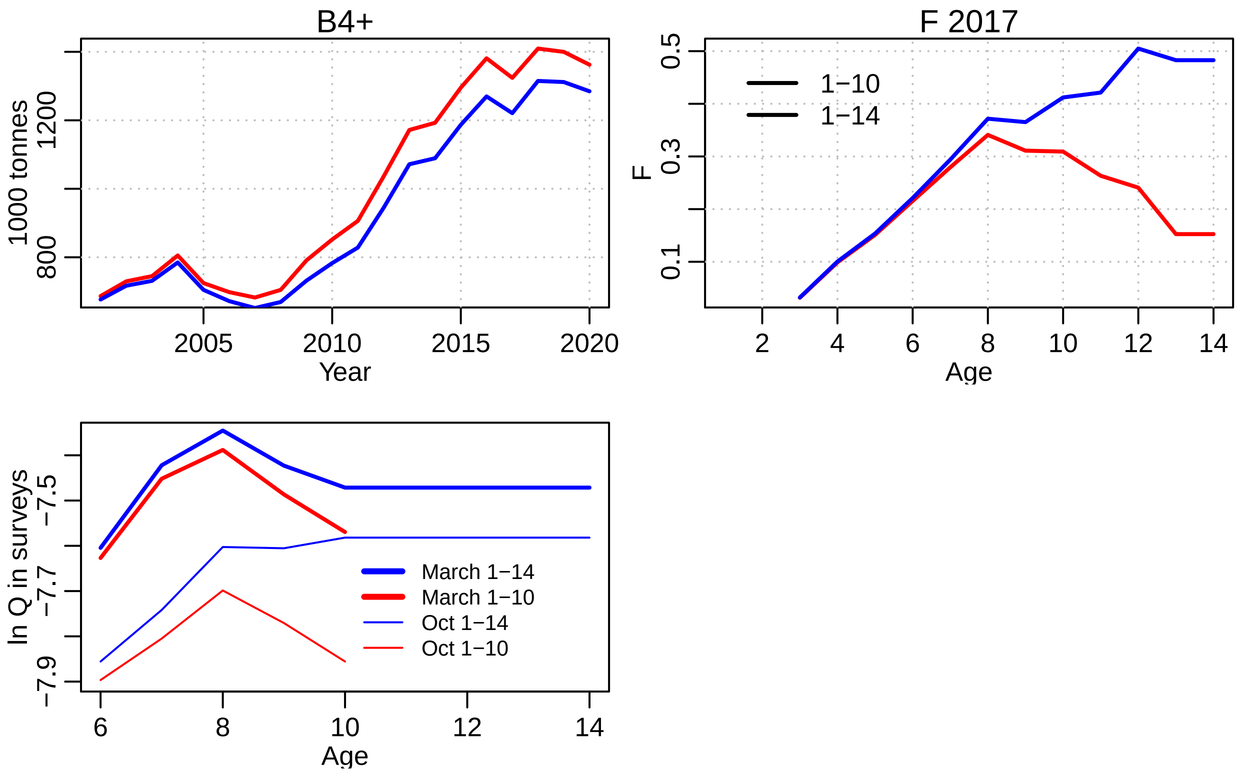


Figure 5: Estimated B4+, fishing mortality by age and Q in the surveys based on tuning with ages 1-14 and 1-10. When 1-14 is used same Q is used for 10-14.

4 Conclusions

1. Of the runs presented model 3 (NEAcod2019block) seems to make most sense. It has the lowest AIC and more sensible grouping of observation variance than the other runs.

2. Canging model 3 so that 5 of the oldest age groups are constrained to have the same F leads to change in log-likelihood from 1181.147 to 1183.934 with 3 parameters less, i.e lower AIC.

3. All data on ages 11-15 have very high observation variance and high correlation leading to random walk taking over the estimation. This should be noticed when a decision of fixing F for 5 oldest vs 2 oldest age groups is taken.

4. Reference points do not have to be changed, historical SSB and therefore B_{lim} are equally wrong or right for all the runs. Changed selection might have small effect on F_{msy} and the management plan but simulating that takes more time than is available here.

5. And then the changes that can not be done this year. Using age 2 and/or 1, extending the survey abundance indices to age 15+, look better at the grouping of the observation variances, combining fleets 15 and fleet 16.

6. If it were not for the increase in F between $2 \times B_{pa}$ and $3 \times B_{pa}$ the advice would be robust to what happens with selection of the oldest age group. The ocean will be full of virtual cod that has no effect on the advice.

4 Conclusions from the reviewers

Paul Regular

"Of the two options, the FINAL run appears to be the better option for providing advice."

Höskuldur Björnsson

"The SPALY run sets the same observation parameters for all age groups. The close to zero values in the past of the older age groups (multinomial errors) make the estimated variances high. But the advantage of the coupling over all the age groups is that the survey indices of the older ages will not be ignored. Therefore, of those two, I would go for SPALY.