WORKSHOP ON THE RE-EVALUATION OF MANAGEMENT PLAN FOR THE ICELANDIC COD STOCK (WKICECOD)

VOLUME 3 | ISSUE 30

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

ICES Scientific Reports

Volume 3 | Issue 30

WORKSHOP ON THE RE-EVALUATION OF MANAGEMENT PLAN FOR THE ICELANDIC COD STOCK (WKICECOD)

Recommended format for purpose of citation:

ICES. 2021. Workshop on the re-evaluation of management plan for the Icelandic cod stock (WKICE-COD).
ICES Scientific Reports. 3:30. 85 pp. https://doi.org/10.17895/ices.pub.7987

Editors

Bjarte Bogstad

Authors

Höskuldur Björnsson • Bjarte Bogstad • Malcolm Haddon • Einar Hjörleifsson •
Margarita Rincón Hidalgo

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

Contents

i Executive summary

WKICECOD 2021 met online to benchmark the assessment methods and evaluate harvest control rules for Icelandic cod. The assessment model was changed during this benchmark. The previous assessment model was a statistical catch at age model with fishing mortality of each age group modeled as a random walk. The new assessment has similar characteristics but is based on having fixed separable periods. The age range of survey indices used in the assessment and the non-linear assumption between stock-in-numbers and survey indices were both extended to include older age groups than before. The effect of these changes resulted in a downward revision of the reference biomass in the last three years compared to the 2020 assessment. The final model has a reasonable retrospective performance albeit with a slight bias in reference biomass estimates (~2.4%).

A Management Strategy Evaluation (MSE) was conducted using a short cut approach to generate assessment and forecast error. The assessment error was auto-correlated to emulate historically observed sequential periods of over- or under-estimation of stock biomass. A stochastic hockey-stick recruitment model was used, assuming that the reduction in productivity estimated in 1985 continues in the future.

The current HCR which uses a harvest rate (HR) of 0.20 of the fishable biomass (age 4+) was found to still be precautionary and to be robust to historical assessment bias. Reference points were re-estimated. The optimum harvest rate in the absence of assessment error is $HR_{msy} = 0.24$ (0.23 including assessment error).

ii Expert group information

Expert group name	Workshop on the re-evaluation of management plan for the Icelandic cod stock (WKICECOD)
Expert group cycle	Annual
Year cycle started	2021
Reporting year in cycle	1/1
Chair	Bjarte Bogstad, Norway
Meeting venues and dates	25 February 2021, online (6-11 participants)
	2-3 March 2021, online (6-11 participants)
	9 March 2021, online (6-11 participants)

1 Introduction

The terms of reference were as follows:

WKICECOD 2021 – Workshop on the re-evaluation of management plan for the Icelandic cod stock 2021/2/FRSG61

The workshop on the re-evaluation of management plan for the Icelandic cod stock (WKICE-COD) will meet online on 25 February 2021 and 2-3 March 2021 chaired by external chair Bjarte Bogstad (Norway), and reviewed by Margarita Rincón Hidalgo (Spain) and Malcolm Haddon (Australia), to evaluate the updated operational assessment model and harvest control rule evaluations for Icelandic cod (cod.27.5a).

The work will be to:

a) Evaluate the appropriateness of data and methods to determine stock status and investigate methods for short term outlook taking agreed or proposed management plans into account for the stocks listed in the text table below. The evaluation shall include consideration of (where applicable): i. Stock identity and migration issues; ii. Life-history data; iii. Fishery-dependent and fishery-independent data; iv. Further inclusion of environmental drivers, multi-species information, and ecosystem impacts for stock dynamics in the assessments and outlook
b) Agree and document the preferred method for evaluating stock status and (where applicable) short term forecast and update the stock annex as appropriate. Knowledge about environmental drivers, including multispecies interactions, and ecosystem impacts should be integrated in the methodology. If no analytical assessment method can be agreed, then an alternative method (the former method, or following the ICES data-limited stock approach) should be put forward;
c) Re-examine and update (if necessary) MSY and PA reference points according to ICES guidelines (see Technical document on reference points);
d) Evaluate the proposed Harvest Control Rule(s) for the Icelandic cod management plan against its objectives and develop conclusions on whether the proposed HCR(s) can be considered precautionary and be used as the basis for ICES fishing opportunity advice for the stock.

Stock leader: cod.27.5a Einar Hjörleifsson

WKICECOD will report by 23 March 2021 for the attention of the Advisory Committee.

The benchmarking of the assessment for this stock is given in Section 3. The management strategy evaluation is described in Section 4. Section 5 contain the references and Section 6 the reviewers' report. The participant list is given in Annex 1 and the request from Iceland to ICES is given in Annex 2, while Annex 3 contains the summary template for Harvest Control Rule modelling. Annex 4 describes an investigation of harvest rates taking cost of fishing and price into account (this was not part of the ToR and was not considered in the review process).

2 Description of the process – request for re-evaluation of the management plan and benchmarking for Icelandic cod (cod.27.5a)

A request for re-evaluation of the management plan and benchmark for Icelandic cod (cod.27.5.a) was sent to ICES from the Icelandic Ministry of Industry and Innovation on 9 November 2020 (Annex 2). Following this, a workshop to answer this request was set up. The background material needed for addressing the request was received from the Marine and Freshwater Research Institute in February 2021.

3 Assessment of cod in 5.a

3.1 Observations

Information on sampling and data compilation are in the stock annex. The method of the compilation of catch-at-age and survey-indices-at-age were last reviewed in the 2015 benchmark and were not revisited in this benchmark. Stock structure was not reviewed during this benchmark.

Below are some selected information and exploration of the input data that may be of current relevance.

3.1.1 Landings by gear

The landings by the principal gears (Figure 3.1) show that the importance of gillnets has been waning, being above 30% in 1982 but less than 10% in recent years. The share of the longline in the catch increased until around 2005, having been around 35% of the total landings since then. The proportional share by trawl has been relatively constant since 2000 being around 40% of the total landings. In the last two years the share of the longline has been decreasing somewhat while the share of the trawls has been increasing.

The proportion landed by age and gear (Figure 3.2) show that gillnets generally target older age groups while the trawl catch somewhat younger fish and longlines even younger age groups. It is of note that since 2001 there has been a gradual shift in the cumulative distribution of landings by age in the two main gears (Figure 3.3), the shift being greater in the trawl gear where there has been a shift of approximately 2 age groups. This is in part a reflection of increased amount/targeting of older fish with decreasing fishing mortality and possibly increasing mesh size by the trawl fleet. Unfortunately the model setup is such that only one commercial fleet is allowed so changes in actual selection/targeting by gear cannot be tested analytically. But it should certainly be explored and may be a recommendation for future work.

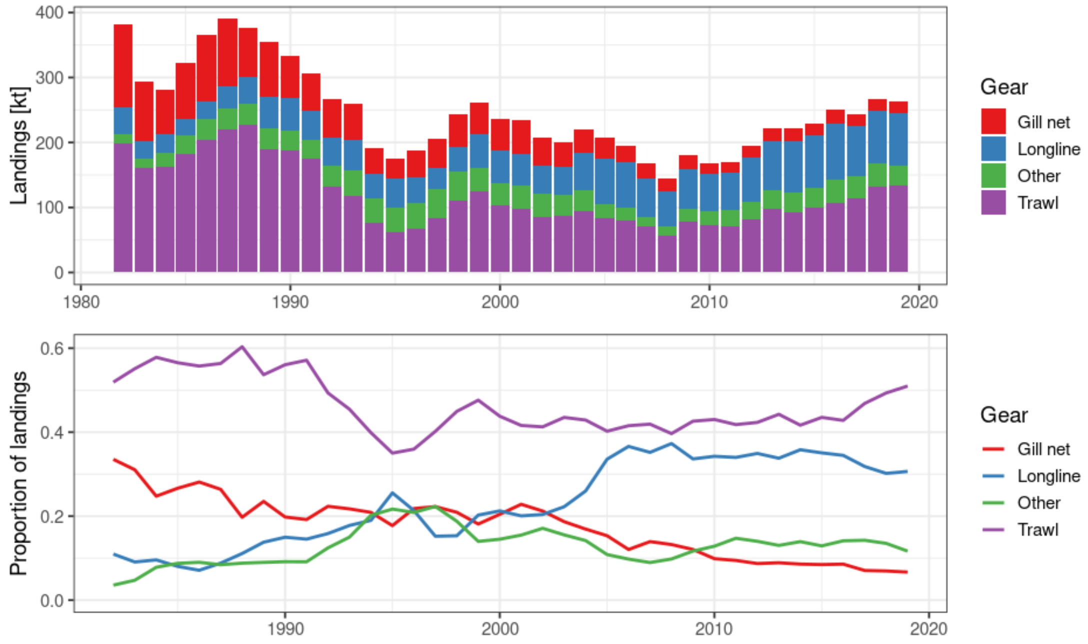

Figure 3.1: Cod landings (upper panel) and proportion (lower panel) by gear.

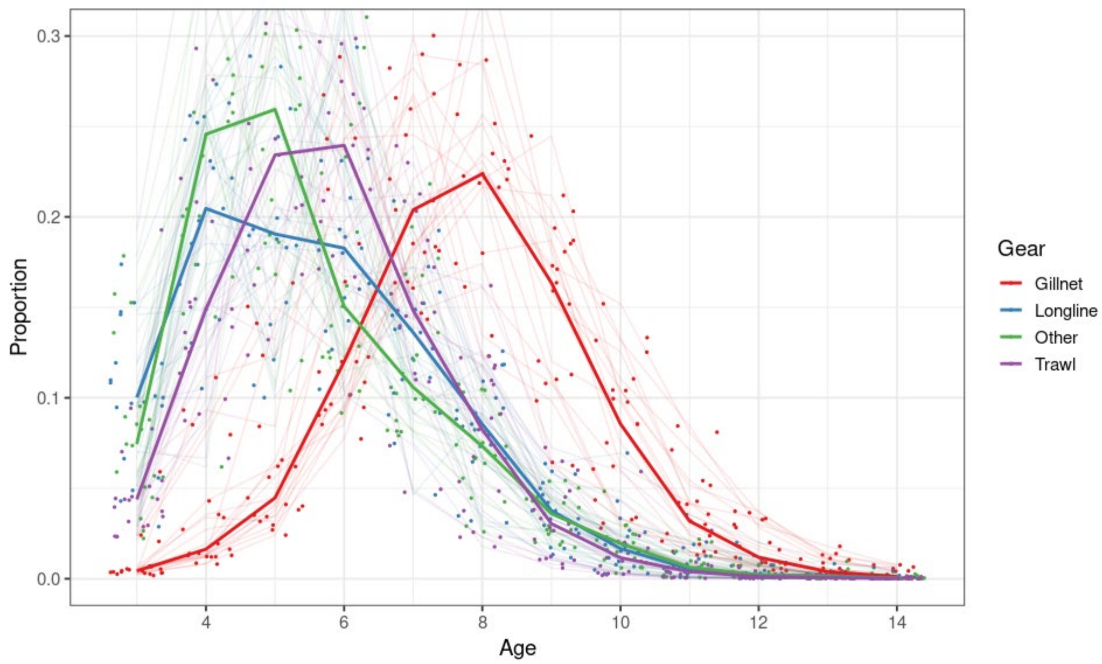

Figure 3.2: Proportional catch by age and gear based on data from 2001 to 2019. Thin line are each year, the thick line the median over years.

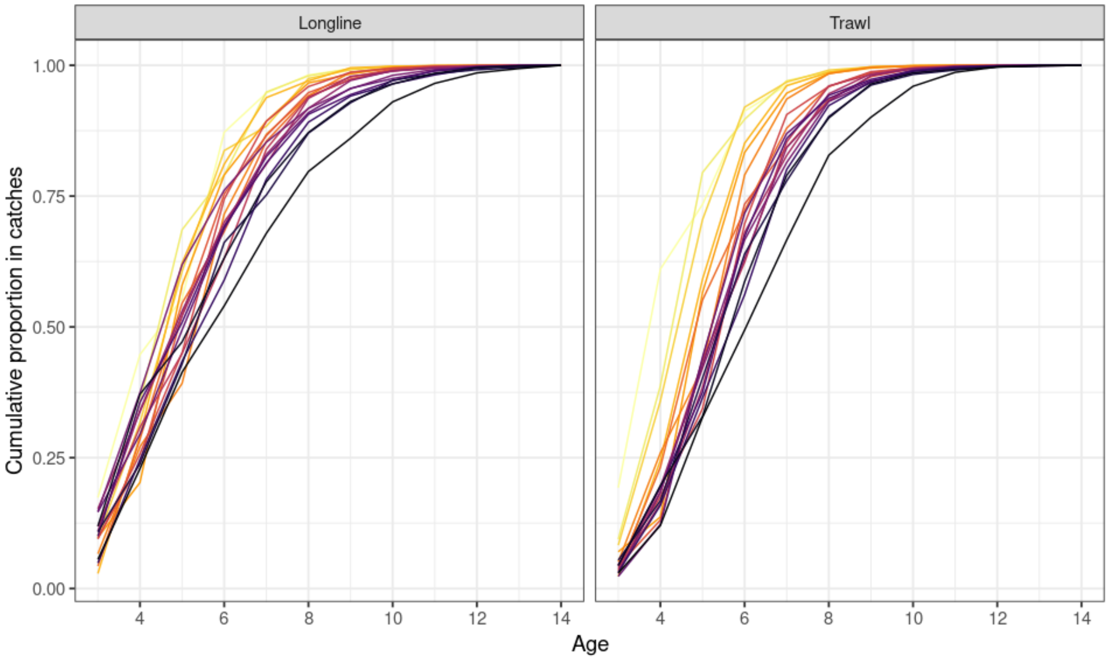

Figure 3.3: Change in cumulative proportion in longlines and trawls catches by age from 2001 (yellow colour) to 2019 (dark colours)

3.1.2 Surveys

Non-commercial input data in the analytical assessments are two annual bottom-trawl surveys, the spring survey (smb) conducted since 1985 and the fall survey (smh) conducted since 1996 (except in year 2011). The spring survey has more stations (~575) and cover the continental shelf proper (< 4-500 m depth) while the fall survey has fewer stations (~375) and also covers deeper waters (<1200 m) (Figure 3.4). Catch rates generally decline at stations deeper than 400 meters (Figure 3.5).

The dynamics of cod in recent years, as expressed in survey biomass indices show a general increase in biomass from around 2010 until the middle of the decade followed by a steep decline in the last 4 years (Figure 3.6). The increase is largely a reflection of reduction in fishing mortality as seen in the substantial increase in larger fish (> 80cm) rather than being a result of significant increase in abundance of younger fish (< 55 cm).

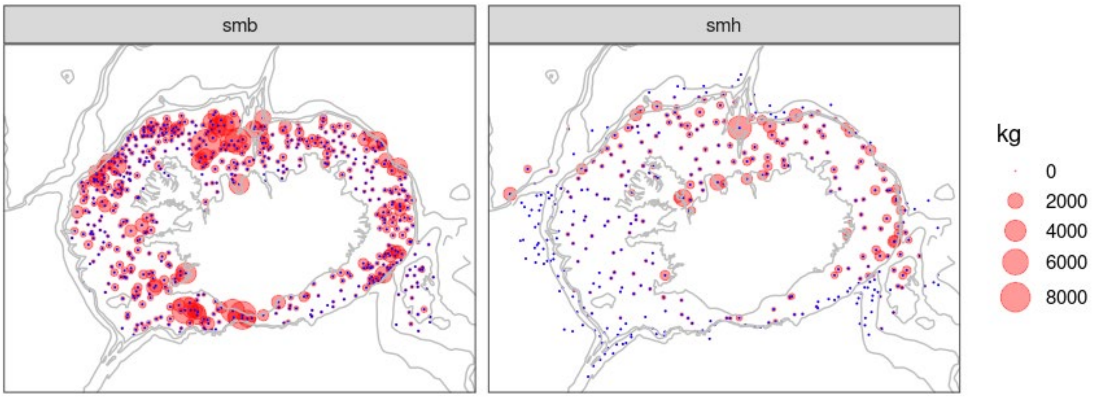

Figure 3.4: Cod catches (kg per hour) in the spring (smb) and fall (smh) surveys in 2019. Blue dots are station location, grey lines 400, 500 and 800 meter contour line.

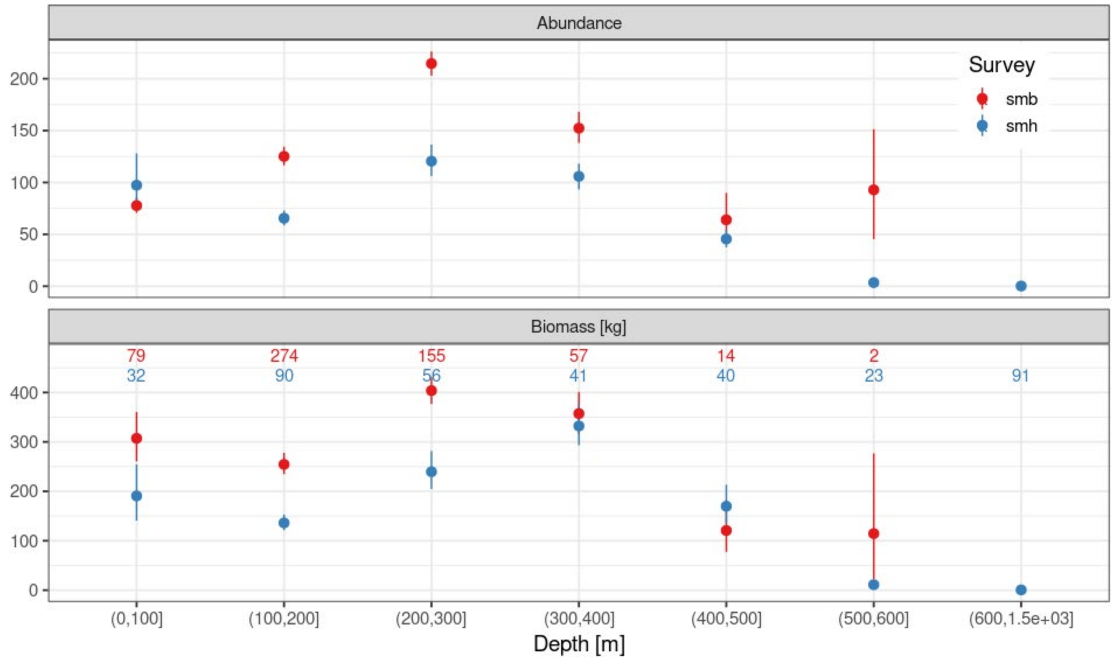

Figure 3.5: Bootstrap mean and standard error of cod catches by depth in the spring (smb) and the fall (smh) surveys. Numerical labels refer to the annual number of stations.

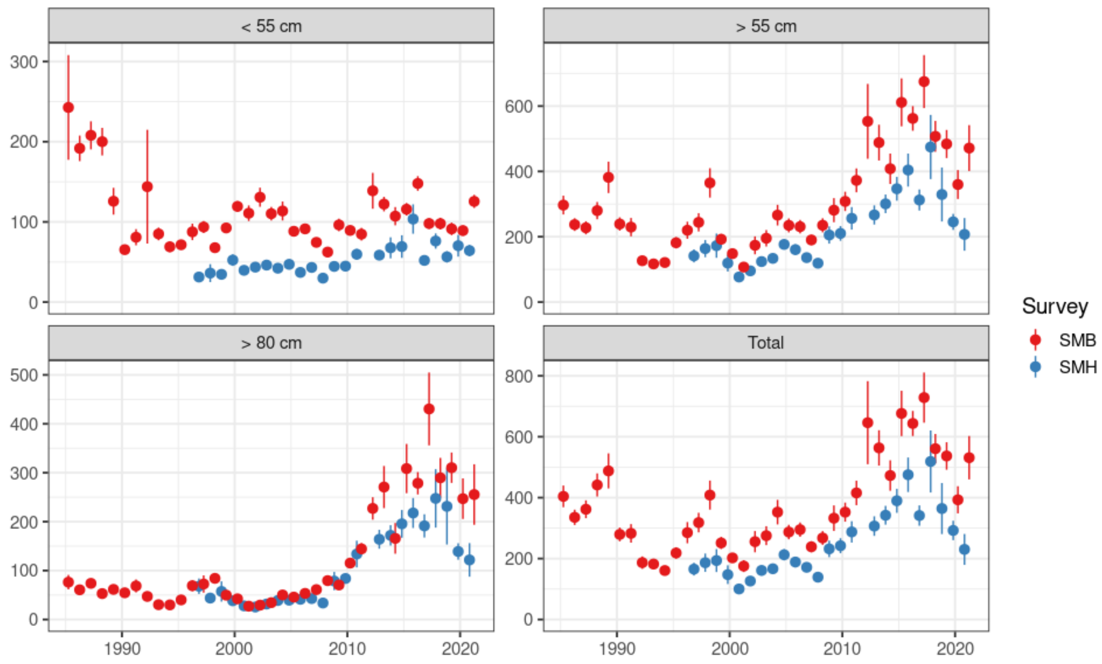

Figure 3.6: Indices of cod in the spring (SMB, red) and fall (SMH, blue) groundfish surveys. Abundance index of fish less than 55 cm, (< 55 cm, top left) and biomass indices of 55 cm and larger (>55 cm, top right), biomass index 80 cm and larger (bottom left) and total biomass (Total, bottom right). The vertical bar shows 1 standard error of the estimate.

3.1.3 Pattern in input data by age

An overview of the catch at age and the survey indices at age (Figure 3.7) show the following overall patterns:

- In period prior to the mid 2000's, catches within a year-class are generally high in the younger age but decline rapidly after age 6 or 7 compared to the more recent period. Similar pattern are observed in the surveys.
- Increase in the catches and the survey indices for ages 8 years and older, the increase in ages 10 years and older being substantial in last decade. The survey indices in these age groups in the start of the time period represented largely noise, while now they can be considered a signal. The increase in older ages in the catches and the surveys occur despite little increase in younger age groups within the same cohort.
- There are indications of improved recruitment in ages 1 in both surveys in the last 10-15 years that are not manifested at later age in the survey nor in the catches.
- A decline in abundance indices of age 5 to 9 years old's in the last 4-5 years.

Any age-based assessment model, in the absence of significant changes in the selection pattern, will interpret the **overall** pattern in the fishable part of the stock (age 4+) as being a decline in total mortality as also manifested by catch curve and ratio analysis (Figures 3.8 and 3.9). Such analysis does though not capture any short-term changes in mortality such, such as the putative increase in recent years.

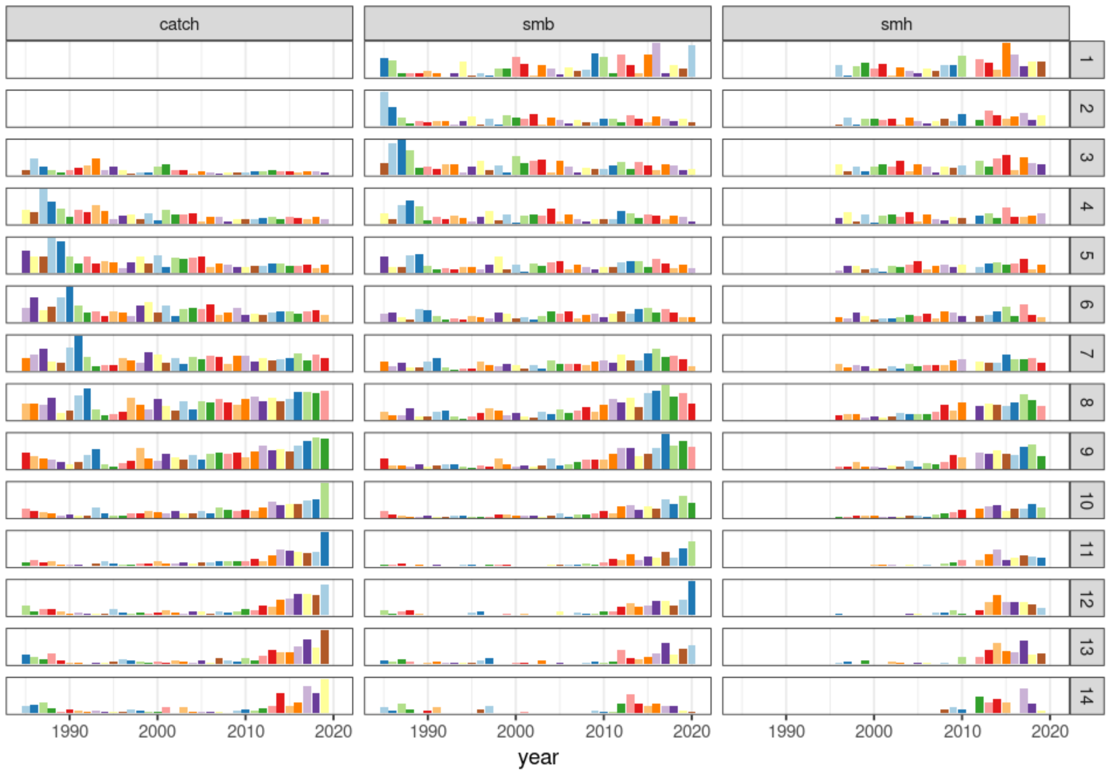

Figure 3.7: Catch at age and age-based abundance indices of cod in the groundfish survey in spring (SMB 1985-2020) and fall (SMH 1996-2019). The indices are standardized relative to the mean over the years 1996-2020 within each age group and variable, colours represent year classes.

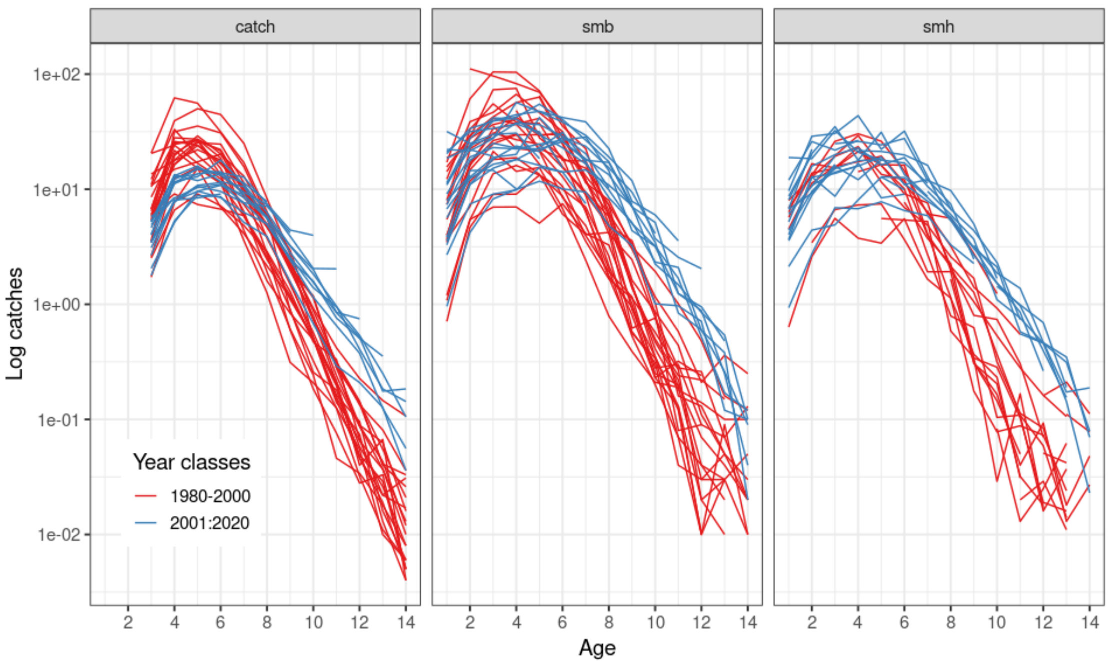

Figure 3.8: Cohort catch curve of commercial catches (catch), spring (smb) and fall (smh) survey indices. Colours indicate periods.

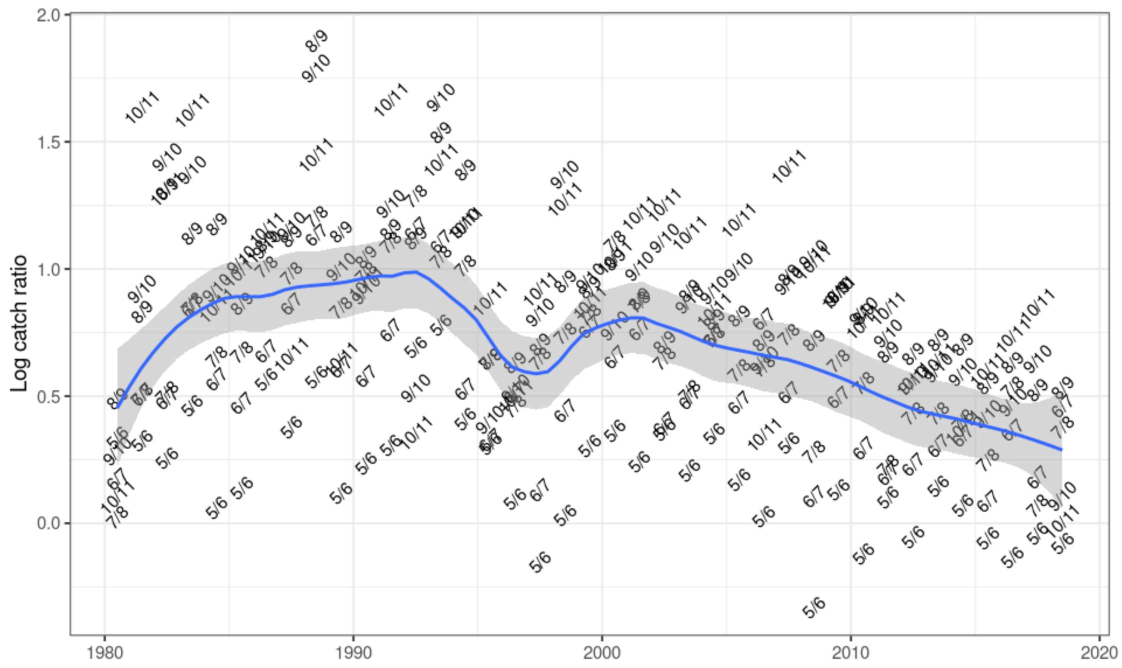

Figure 3.9: Log ratio of catches for age groups 5 to 11.

Since cod do not start to enter the fishery until age 3 and 4, any trend in the log catch ratios in survey indices younger than age 3 can not be assigned to changes in fishing mortality. Log catch-ratios between age groups within a cohort (Figure 3.10 and 3.11) show the following patterns:

- There is an indication of increase in log-ratio in the pre-recruit age classes 1 vs 2 and 2 vs 3 in both the smb (spring) and smh (fall) surveys. The increase in age 1 vs 2 in spring (smb) is higher than in the fall (smh) survey, here accounted for by including a change multiplier in q at age 1 in year 2003 for the spring survey. The changes are either a signal in change in mortality, catchability or a change in the distribution of young fish over time relative to survey coverage (younger fish are generally thought to be in more shallow waters).
- The slopes of the log catch ratio in age 3 vs 4 in the surveys are not significantly different from zero.
- Increase in the log-ratio in the catches of age 3 vs 4 and 4 vs 5 in the most recent years may be an indicator of increased mortality.
- The declines in the slope of the survey indices for older age groups that have entered the fisheries are largely a reflection of changes in fishing mortality in recent decades and are generally similar in magnitude to those observed in the catches.

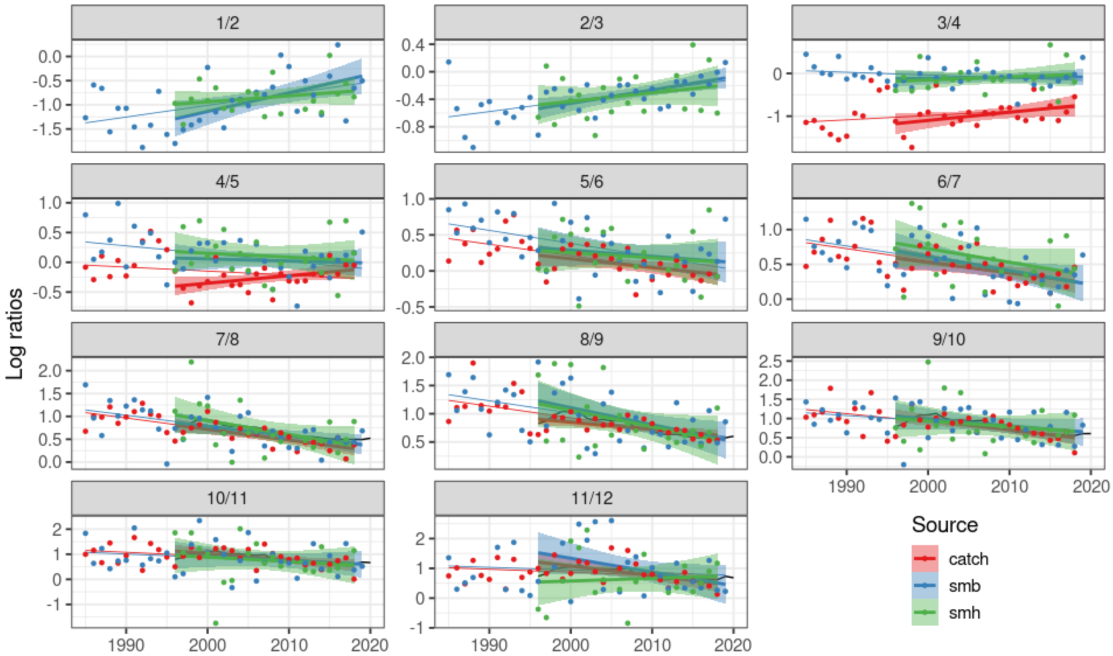

Figure 3.10: Log catch-ratio of the catches, spring (smb) and fall (smh) surveys between age groups (panels). Linear trend lines are shown for the period 1996 to 2020, with a thin line showing the trend line for the spring survey (smb) over the whole time period (1985-2020).

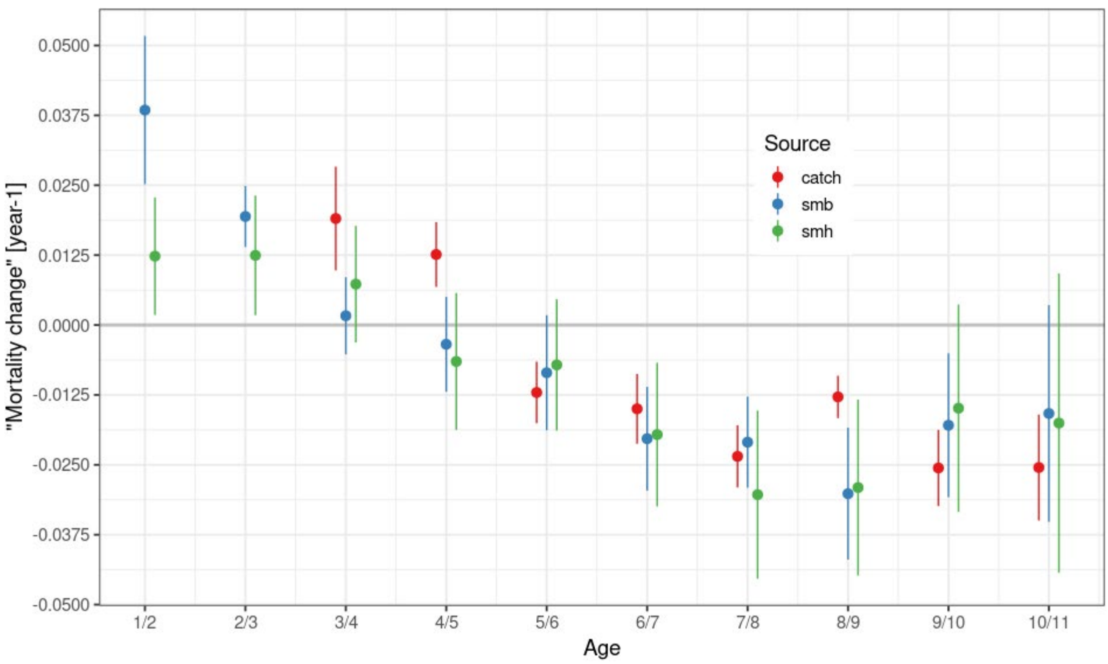

Figure 3.11: Slope of a linear trend of log-ratios from 1996 onwards of catches, spring (smb) and fall (smh) surveys.

The most recent points in the age 2-3 relationship (Figure 3.12) are below the predicted average and the same applies to age 1 vs age 2. For ages 3-6 in 2019 the predicted survey index in 2020 is on the average line but for those age groups many of the more preceding 2-4 years are above the average line.

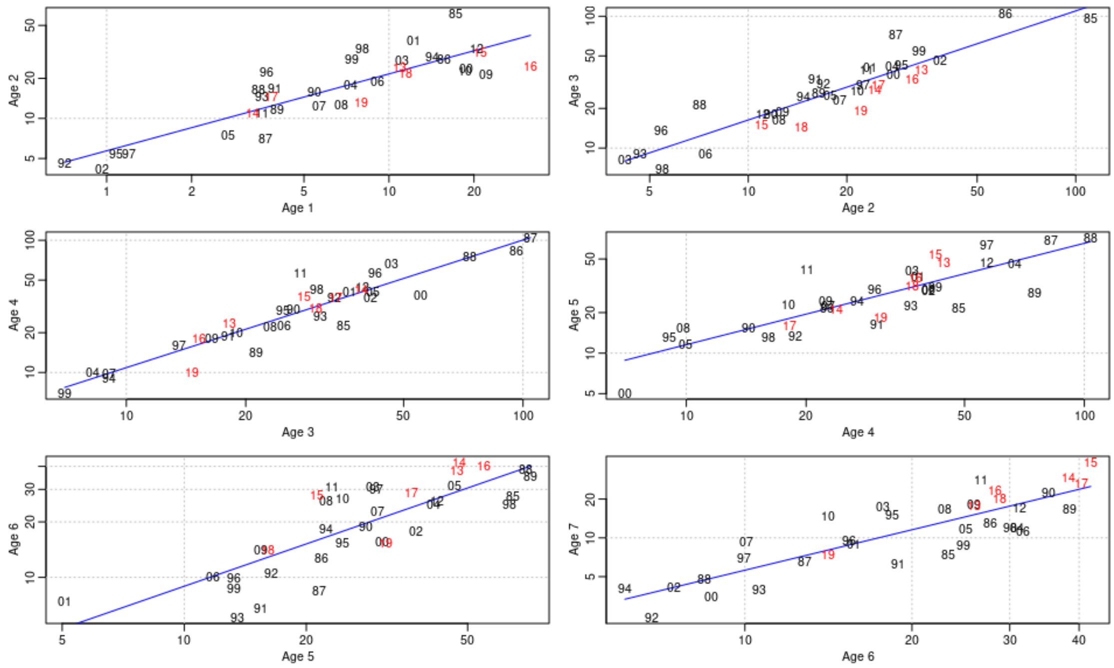

Figure 3.12: SMB index at age A+1 against index of the same year class one year earlier. Line fitted to the pairs shown in the figure.

3.2 Analytical assessments

The key stock metrics based on the 2020 assessment (Figure 3.13) indicate the following main pattern:

- Biomass is high in the beginning of the time series in part as a result of low fishing mortality (was also low during WWII and the years thereafter). The 1945 year class is estimated to account for at least 500 kt of the initial biomass but a large part of year class 1945 migrated from Greenland in 1953 (based on extended catch at age not shown). Migrations from Greenland have been estimated in 1958, 1959, 1960 and 1962 although the estimates of those migrations are confounded with estimates of initial numbers-at-age.
- Despite good recruitment, there is a general decline in biomass until around the 1990's. Catches were also declining, but at a lower rate mostly driven by increase in fishing mortality.
- The stock remained low until the late 2000's, but has increased in the last decade largely driven by reductions in fishing mortality rather than significant improvement in recruitment.

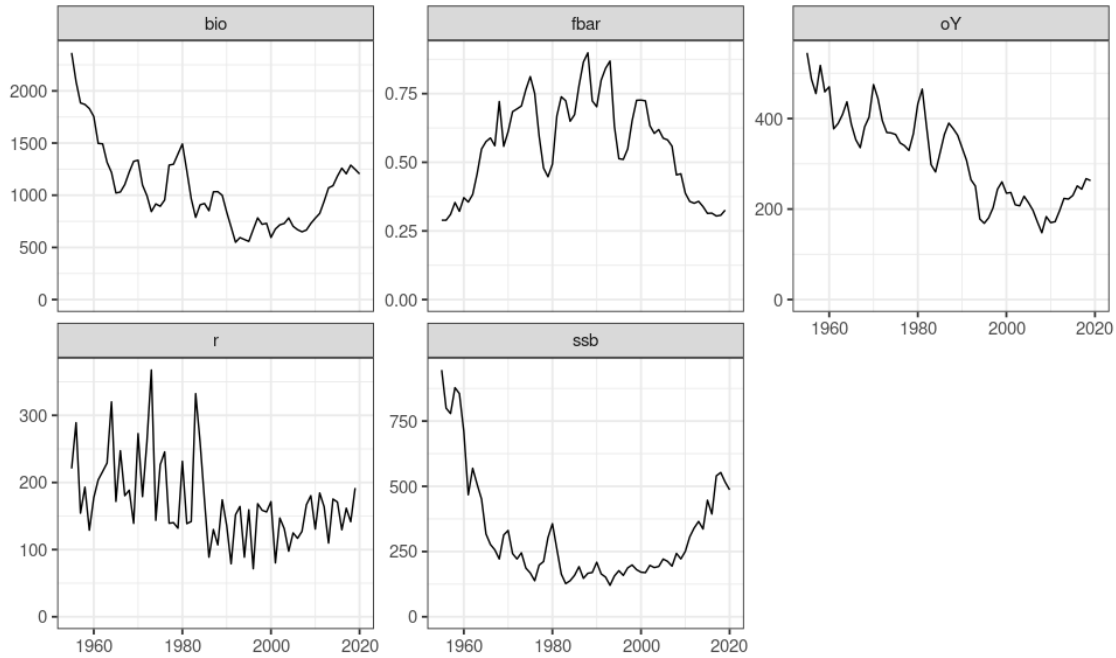

Figure 3.13: Reference biomass (bio, kt), fishing mortality (fbar, mean ages 5-10), yield (oY, kt), recruitment (r, millions scaled to age 3, x-axis refers to year class) and spawning stock biomass (ssb, kt) from the 2020 assessment.

3.2.1 The 2019 and 2020 assessments - issues

The last benchmark for this stock was in 2015. The 2020 assessment was done by the Marine and Freshwater Research Institute, Iceland and changes in the assessment setup in 2020 (described below) was not formally reviewed by ICES.

The annual advice and subsequently the TAC decision for Icelandic cod has, since 2002, been based on framework often referred to as adcam (admodel-builder code, see more on that later in this document). Up to and including the 2019 assessment, only survey age groups 1 to 10 have been used in the assessment, first only using the spring survey (smb), and then the fall survey (smh) that was included since 2010. A non-linearity between stock-in-numbers and survey indices was assumed for ages 1 to 5 for the spring survey and ages 1 to 4 for the fall survey.

In the recent assessment using only these ages in the tuning has resulted in temporary increased dome-shaped selection pattern (Figure 3.14). While this may have not been an issue while catches were relatively low in older age groups it currently results in the generation of substantial amount older fish in terms of stock in numbers. With respect to advised catches this may not be a problem using conventional ICES methods (catch advise based using recent estimated selection pattern). However, this will result in higher catches in knife edge decision rule such as used with Icelandic cod, where the advised catch is in principle 20% of the reference biomass (B_{4+}) at the beginning of the assessment year. A dome-shaped selection pattern will in both cases affect the estimates of the "apparent" SSB, where the estimated selection pattern is not taken into account, only maturity. Dome-shaped selectivity is a real possibility with Icelandic cod, the gillnet fisheries that were targeting spawning cod are at an historically low percent of the total and large/older fish might migrate off the continental shelf outside the spawning season. However, assuming a dome-shaped selectivity as the null hypothesis, with implications regarding biomass, warrants further research.

In the 2020 assessment cycle one more year of declines in survey indices (see figures 3.7 and 3.15) were observed, with the spring survey in 2020 being particularly low. These declines were not manifested in the predicted biomass from the analytical model, but it was clear that the predicted biomass was biased high. A limited exploration with respect to the dome shaped selection pattern was performed in the 2020 assessment cycle resulting in inclusion of the older ages (11-14) into the assessment tuning, constraining ages 10 and older to have the same **q**. This resulted in a selection pattern that approached a maximum at older ages and lower absolute biomass estimates in the last decade (Figure 3.14).

The model seems to capture the survey indices of the older age groups reasonably well (Figure 3.16). The trendlines fitted to the data have a slope of 0.97 for the spring and 1.11 for fall survey, not indicating any major non-linearity. However, it did not capture better the contrast in the survey dynamics in the last decade (Figure 3.15). It is noted upfront that none of the model setups explored during the benchmark did (see e.g. Figure 3.24).

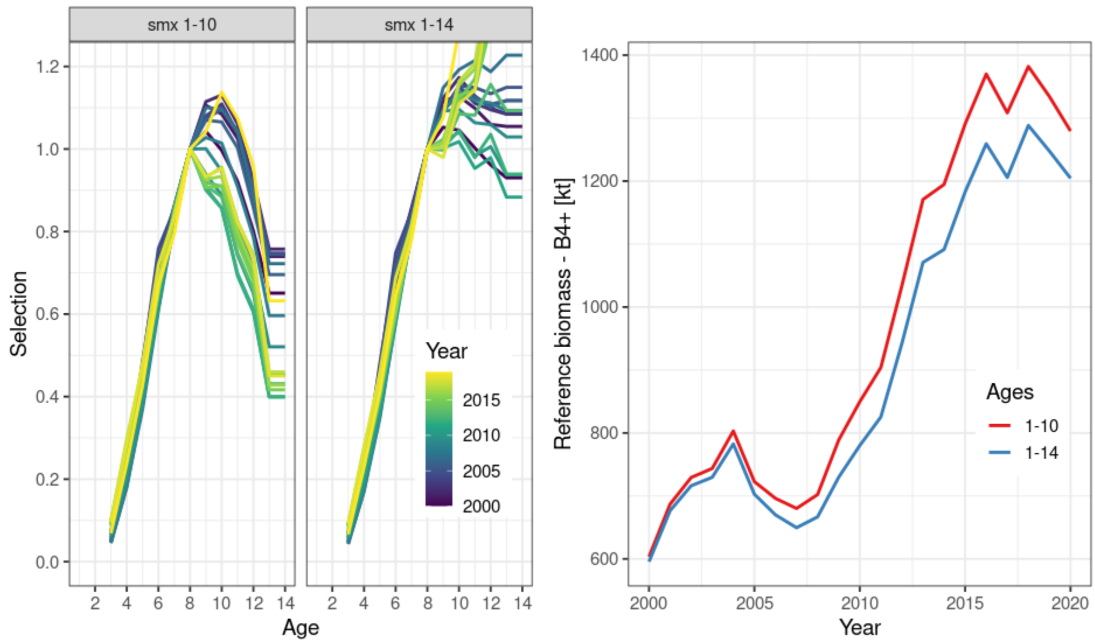

Figure 3.14: The effect of inclusion of older age groups (11-14) in the tuning on selection pattern (here scaled to age 8) and reference biomass estimates in the 2020 assessment based on adcam-model using both surveys in the tuning (smx 1-10: age groups 1 to 10, smx 1-14: age groups 1 to 14).

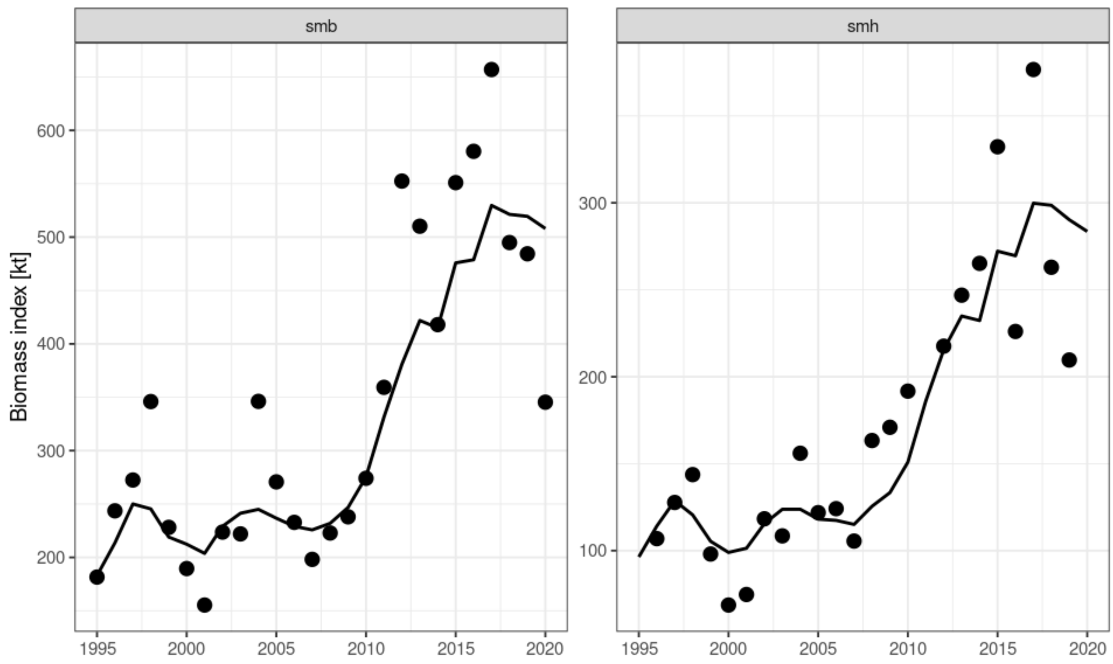

Figure 3.15: Sum of observed (points) and predicted (line) age-based indices, for the spring (smb) and fall (smh) expressed as biomass from the 2020 assessment (adcam model, both surveys used, ages 1-14).

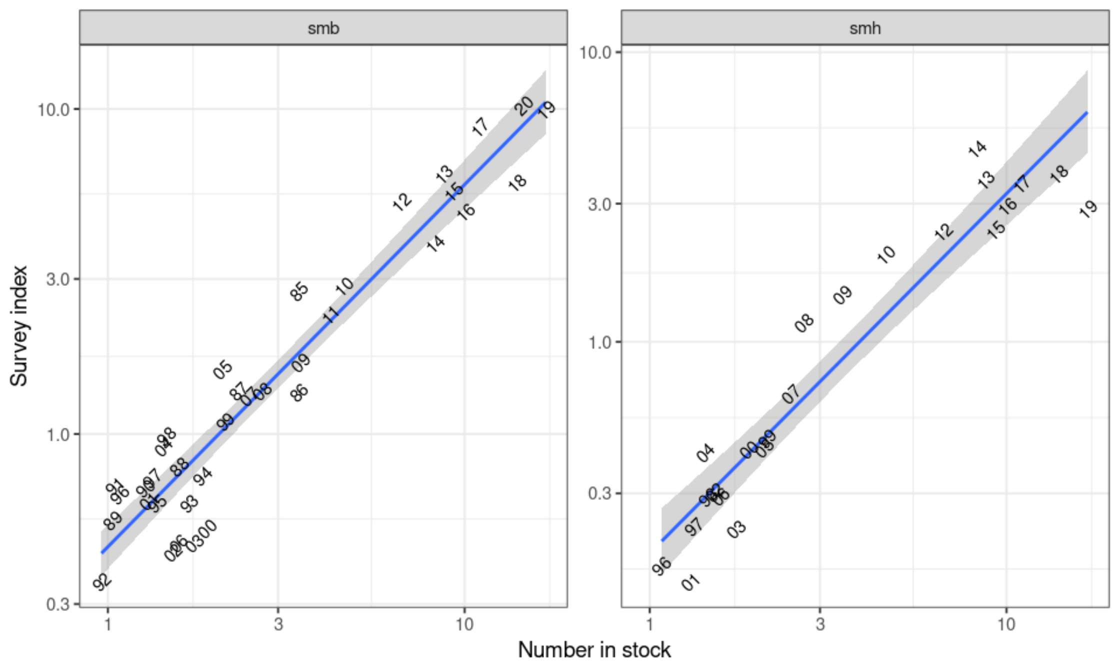

Figure 3.16: Stock in numbers of ages 10 years and older vs survey indices (smb: spring survey, smh: fall survey) for the same age range based on the adcam tuning with ages 1 to 14. The labels refer to year.

Further analysis showed that the inclusion of older age groups resulted in an increase in systematic retrospective pattern that was not as apparent when tuning only with age groups 1 to 10 (Figure 3.17). A systematic bias like that observed may warrant an automatic rejection of the inclusion of older age groups in the tuning. However, it should be kept in mind that the range in the observed survey indices in older age groups has been expanding in the last decade (Figure

3.7) and thus the estimate of the catchability in these age groups is improving with each year added.

Some notes on the calculation of the bias may here be warranted. Firstly, it is not exactly the same as Mohn's rho statistics conventionally now used in ICES, but on the log-ratio ($bias = ln(By, y/B_{y,term})$, in-house referred to as Ziggy's rho, Jonsson and Hjorleifsson (2000)). Secondly, in the calculation of mean bias we prefer to compile the measure based on peels that do not include the most recent assessments (as is now conventional in ICES) but is based on some earlier period. Ideally this should be the period where the assessment has "converged." But due to slow convergence in the model this is not possible, at least when using both surveys (earliest retrospective peel is the 2002 assessment when we only have 6 years of data for the fall survey (smh)). In the case of the Icelandic cod, the bias calculated based on the ICES conventional period (last 5 years peel) is only 2% (tuning with survey ages 1 to 10) and 4% (tuning with survey ages 1 to 14) while taking the peels from years 2002 to 2015 gives the equivalent metric of 4% and 9% (Figure 3.17). Whatever the case it is also recognized here that one cannot really distinguish between assessment bias and autocorrelation given the time periods considered here.

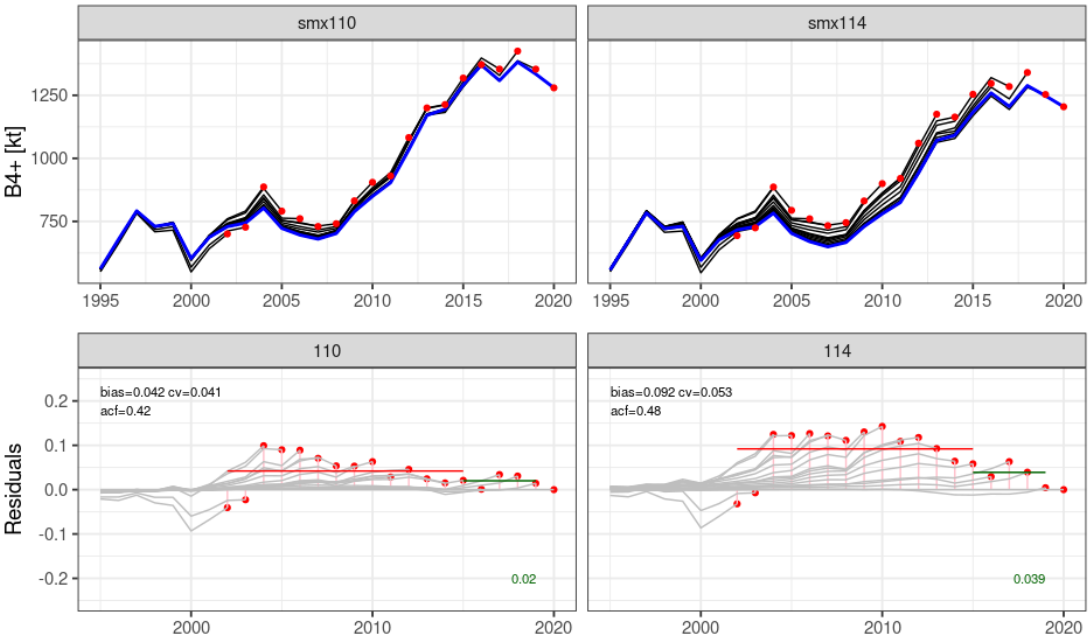

Figure 3.17: 2020 assessment. Comparison of retrospective pattern in the reference biomass (B4+) when tuning with age groups 1 to 10 vs. age groups 1 to 14. Upper panel: conventional view, 2020 assessment (blue line), earlier peels (black lines, red dot showing terminal value. Lower panel: Retrospective patterns expressed relative to 2020 assessment. Statistics based on assessment year peels 2002 to 2014, additional bias measure for the last 5 peels also expressed.

Over the last 10 assessment years, tuning separately with the spring (smb) and fall survey (smh) has resulted in lower estimates compared with tuning with the spring survey only vs. the fall survey only, with fishing pressure showing a reverse difference. The difference in the reference biomass in the terminal year between those tunings has generally been around 20% with the tuning using both surveys (smx) normally being in the middle (Figure 3.18).

Figure 3.18: Comparison of reference biomass (B4+) and fishing mortality (average age 5-10) estimates based on tuning with the spring survey (smb), fall survey (smh) and both surveys (smx). Based on adcam 2020 assessment settings.

3.2.2 Updating the cv by age-profile and the autocorrelation model

In the forward based model framework used here (both adcam and muppet - see later) the CV for each age groups can-not be estimated internally in the model. Instead the standard deviation for each survey and age group is estimated using a VPA framework where the starting stock-in-numbers of the oldest age groups are those obtained from the forward based framework. This profile is then used as input in forward models and only an overall weight for each survey is estimated internally.

Last time the profile was updated, in the 2015 benchmark, the survey model used was exactly the same as in the assessment (including the power estimated on age 1-5 in smb, 1-4 in smh). This profile resulted in age 2 in the spring survey having relatively more weight in the assessment as age 2 indices fitted well within the converged assessment and the same applied to age 1 and age 2 in the autumn survey (Figure 3.19, upper panels).

In the 2021 benchmark the updated CV-profile was based on using a linear relationship for all age groups. The idea is to give lower weight to age groups where the power is high, but the reasons for there being a nonlinear relationship are poorly understood. Unless specified otherwise, an updated age-profile is used in subsequent analysis.

The updated profile results in age groups 1 to 5 data getting a relatively lower weight in the assessment compared with age groups 6 to 9 (Figure 3.19, lower panels). For the youngest ages the updating resulted in a relatively greater increase in CV in the fall than the spring survey. The overall implication in the updating of the profile is that the "earliest data" from each cohort receives lower weight in the terminal year of the assessment than the more recent ones. In addition, the influence of the fall survey on the estimates of the incoming recruiting ages becomes relatively less. Updating the age-profile every assessment year rather than at each benchmark may be the recommended procedure in the future.

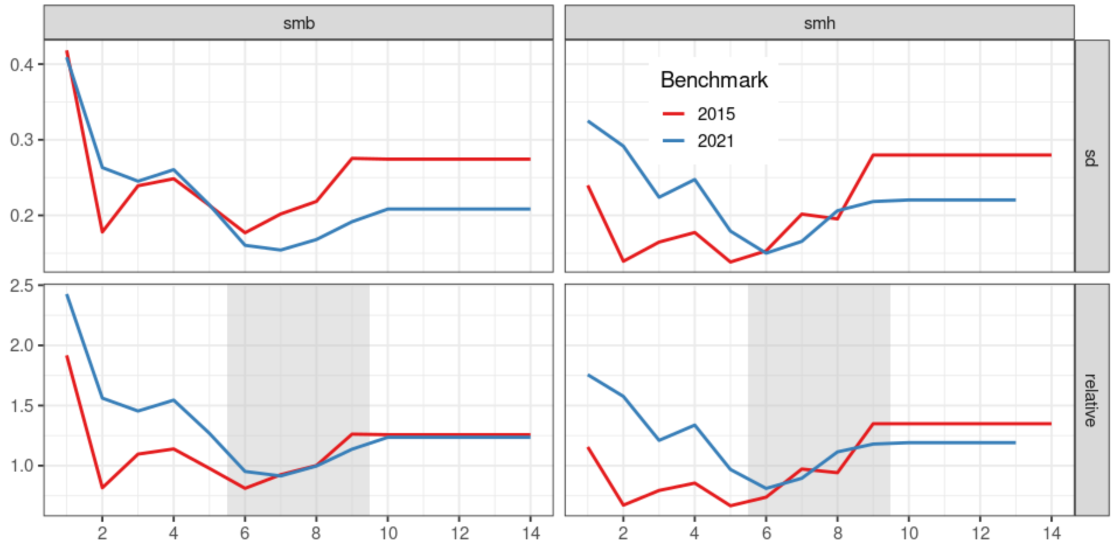

Figure 3.19: Comparison of the estimated standard deviation from the 2015 benchmark and the updated estimates for the spring (smb) and the fall (smh) survey indices. Upper panels: Calculated values, lower panels: Values scaled to the mean of age 5-9.

In the 2015 benchmark setup, age-groups 1 and 2 were assumed to be uncorrelated with the older ages. In addition, the "distance" between age-groups was nonlinear function of age so the "distance" e.g. between ages 10 and 9 is less than between ages 4 and 3. The correlation model was updated so that the correlation matrix was the same for all age groups:

$$Corr_{a1,a2} = \rho^{abs(a1-a2)}$$

The effect of these changes resulted in a downward revision of the reference biomass in the last three years compared to the 2020 assessment (Figure 3.20), where the terminal estimate of the reference biomass changed from 1205 kt (the 2020 assessment) to 1171 kt with updated survey standard deviation-profile ("Weights"), and then further to 1121 kt when changing the age correlation model. The effect of considerably reducing the weight of the random walk fishing mortality parameter ("Reduced walk") had only a minor effect. Further analysis of the effect of correlation indicated that the inclusion of ages 1 and 2 in the correlation accounted for most of the effect.

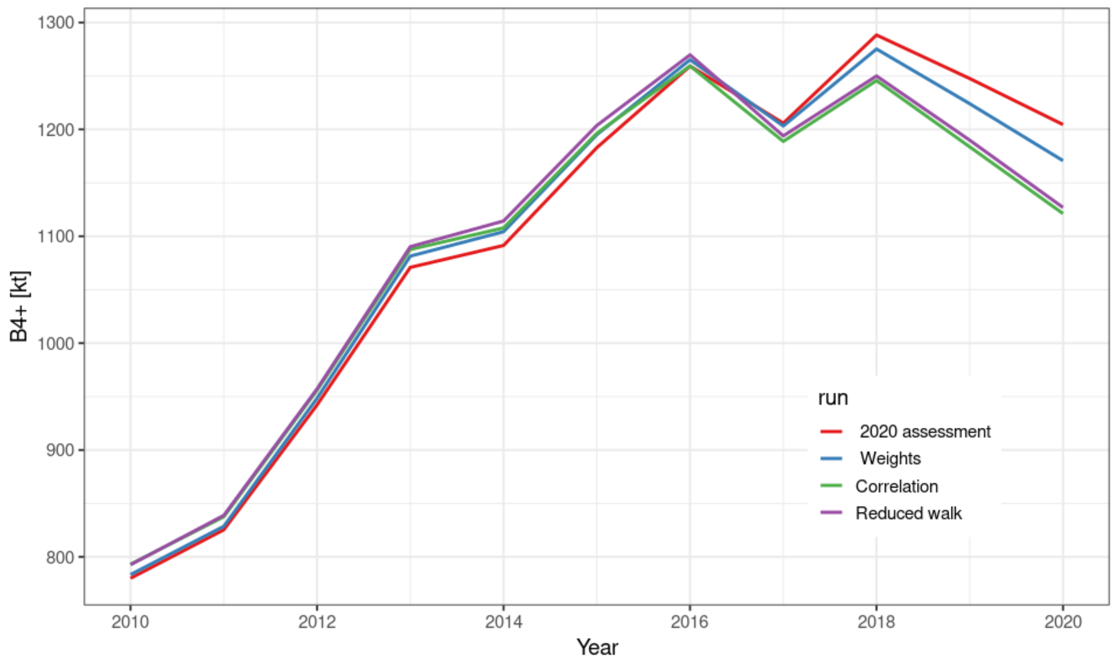

Figure 3.20: Effect of updating the CV age-profile and correlation model on the adcam 2020 assessment reference biomass (B4+). See further explanation in the text.

3.2.3 adcam vs. muppet

The annual advice and decision on the TAC for cod has since 2001 been based on a model referred to as adcam. Harvest control rule evaluation has however been performed in a framework that is now referred to as muppet and is a program that has been used for assessment and HCR evaluations for a number of stocks.

Both the ADCAM and muppet model work similarly as assessment and short term prediction tools except for modelling of fishing mortality that is just a separable model in muppet but the whole F_{ay} matrix is estimated in ADCAM, using interannual constraints (multivariate random walk). A comparison of the "updated" adcam 2020 assessment and the equivalent muppet run show very consistent results (Figure 3.21) especially with respect to biomass and recruitment, except for the start of the time series of biomass. Since the difference in key metrics between the two modules are minor it was decided to switch to the muppet module as a basis for annual advice for cod for the following additional reasons:

- Difficult to maintain two admodel-builder codes.
- The muppet code in cleaner and neater.
- The input files are easier to generate for muppet.
- The separable model has the forward hcr-simulation inbuilt.
- Development in progress for an assessment in the loop in muppet.
- Muppet is used in the Icelandic haddock and saithe assessments.

Hence the adcam module is not considered further in subsequent explorations and modelling gradual change in fishing mortality is much better done using a random-effects model.

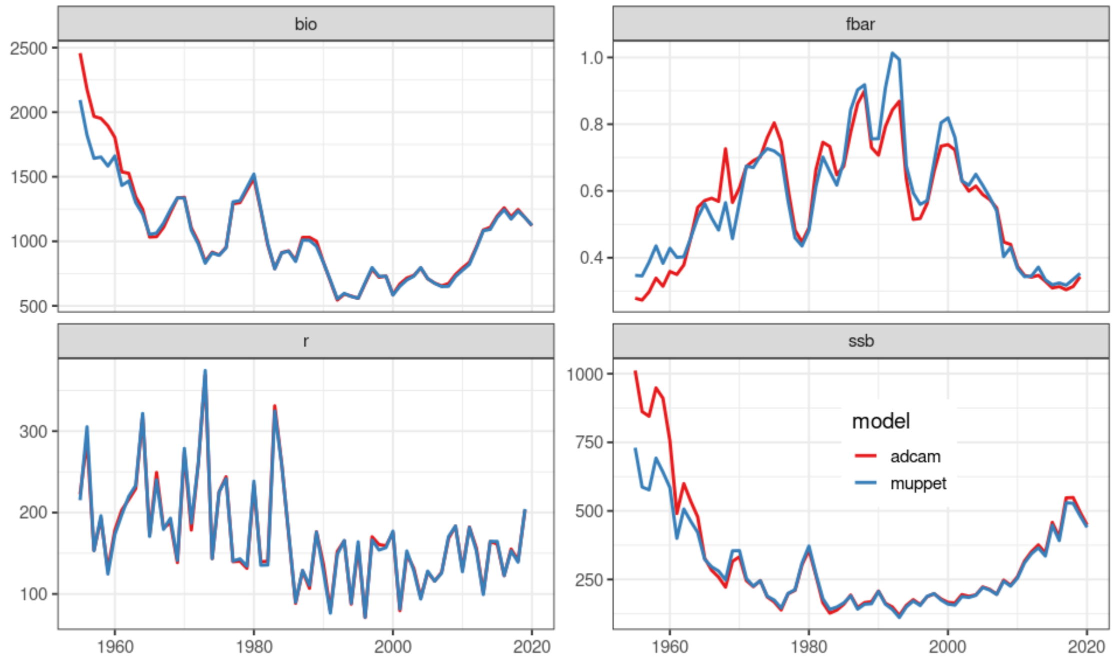

Figure 3.21: Comparison of key assessment metrics of the adcam and muppet module (later referred to as run '1') using the updated 2020 assessment tuned with both surveys, ages 1 to 14. bio: reference biomass (B4+, kt), fbar: average fishing mortality ages 5 to 10, r: recruitment scaled to size at age 3 [millions], ssb: spawning stock biomass [kt].

In the muppet model setup there are four separable periods specified, 1955-1975, 1976-1994, 1995-2007 and 2008 on-wards. A comparison of the estimated selection pattern for four models, vpa, adcam, sam and muppet (Figures 3.22 and 3.23) show that the selection pattern in younger fish relative to selection at age 8 are very similar for vpa, adcam and sam. The selection pattern in muppet, being fixed are in principle the average selection within each selection period. Within each separable period one observes some deviations pattern with time relative to the more flexible models. Of most interest is probably the last period where there are indications of increased targeting of younger fish relative to age 8. The question is always if these changes are transient or permanent, something that cannot be answered at present. The assumption of the mean selection pattern assumed in the forward simulations for the HCR evaluation are likely of most concern. Earlier analysis based on saithe (has the same catch rule as cod) showed that the form of the decision rule is relatively robust to assumed future selection pattern (unless it is dome shaped). An attempt will be made to demonstrate this in the current simulations.

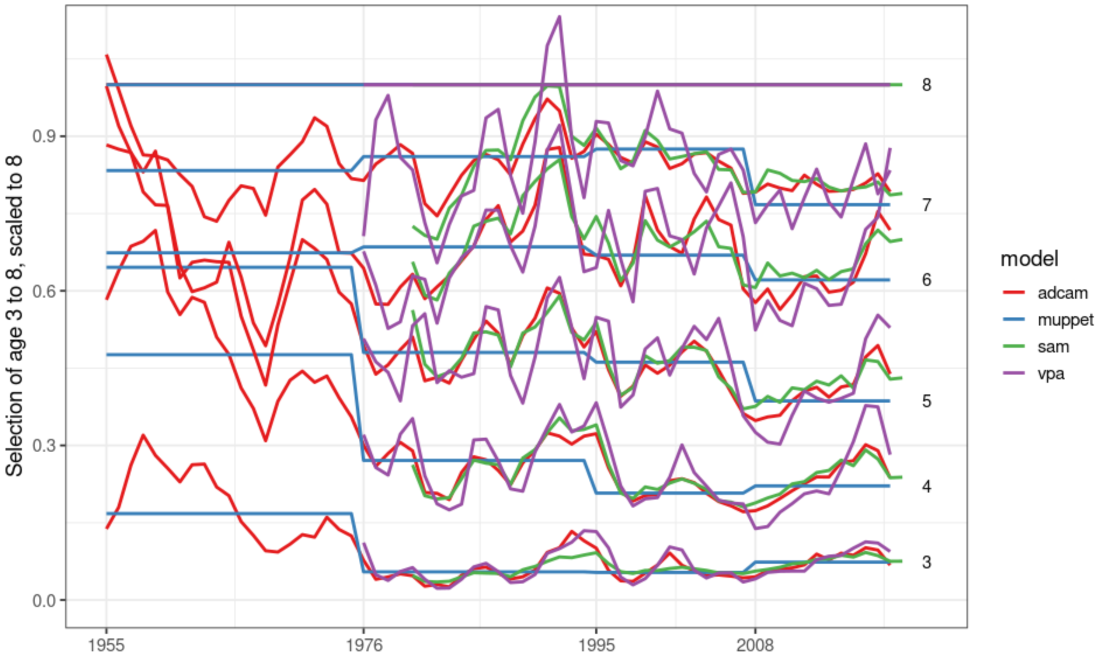

Figure 3.22: Selection pattern of adcam, muppet, vpa and sam for ages 3 to 8, scaled to age 8 (runs based on smx1p setting, a detailed description on this implementation is provided in the next section). Selection pattern from vpa prior to 1976 not shown because catch at age is quite noisy. sam assessment is only run from 1980. The years indicated on the x-axis is the starting year for each separable period used in muppet.

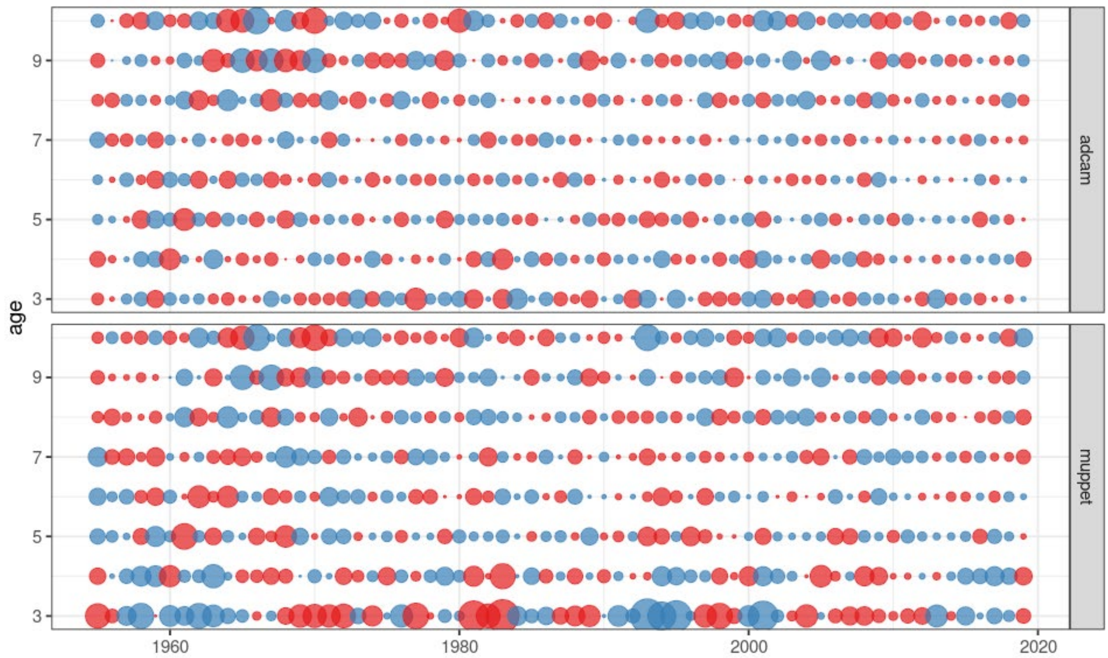

Figure 3.23: Raw catch residuals for ages 2 to 10 from the adcam and muppet (runs based on smx1p setting, a detailed description on this implementation is provided in the next section). Red: negative, blue: positive residuals.

3.2.4 Preamble and overview to subsequent analysis

3.2.4.1 Preamble

During this benchmark process the main focus has been on investigating the retrospective pattern that are observed when tuning with both surveys using age groups 1 to 14 as well as investigating the lack of response of the predicted reference biomass, given the decline in the survey indices in recent years. There is some conflict in those approaches as a model responding more slowly may get less of a retrospective pattern, but it has not yet responded as expected, and too much inertia in models is not desirable when there is a stabiliser in the HCR.

The approach taken was to run a sweep of setups:

- Run most data/model setup on the spring survey (smb), the fall survey (smh) and both surveys (smx).
- Explore the effect of extending the power model to age 9.
- Given that year classes last now for long (10 years or more) in both the surveys and the catches and because of trends in the signals in the youngest ages groups (ages 1 and 2) explore the effect of removing these youngest ages.
- Explore the effect of including the youngest age group (ages 1 and 2) as a separate fleet but giving them relatively low weight.

In subsequent text the different model setups are labeled as follows:

- **sm*1**: Same as the 2020 assessment setup: Survey age groups 1-14 in spring and 1-13 in fall, non-linearity on ages 1 to 5 in spring and 1 to 4 in fall.
- **sm*1p**: Same as **sm*1** but non-linearity assumed for all age groups up to 9
  - **smx1_3p**: A variant of **smx1p** but only using ages 3 to 13 in fall survey - only run with both surveys.
- **sm*3**: Same as **sm*1** but only using ages 3 years and older.
- **sm*3p**: Same as **sm*3** but non-linearity assumed for age groups 1 to 9
- **sm*3pa**: Same as **sm*3p**, but single non-linearity parameter estimated for age groups older than 5.
- **smx3parw12** Same as **smx3pa** but survey indices 1 and 2 added, given externally a low weight.

When tuning with the spring survey only, the acronym used in the document is **smb***, fall survey only is **smh*** and both surveys is **smx***.

In runs referred to as **sm*3pa** the observation model was changed from a conventional power relationship for ages 6-9 in the spring survey and 5 to 9 in the fall (i.e. only power on ages 3-5 and 3-4) to a model working on total survey numbers by including an extra term in the equation:

$$log(\widehat{U_{ya}}) = log(q_a) + p_a \times log(N_{ya}) + log(\delta_y)$$

$\delta_y = \frac{(\sum_a \widehat{U_{ya}})^\gamma}{\sum_a \widehat{U_{ya}}}$ Equation 1.

$delta_y$ is a yearfactor for the survey.

where γ is estmated parameters for each survey. The powers p_a for younger fish reflect distribution in shallow waters and are estimated also for ages 3-5 as was usually done earlier.

What this model really does is to compile the total number of fish in the normal way, then scaling all numbers up by putting power on the total number. Tuning is then done in the same way as before based on observed and predicted survey indices by age. For each survey 1 parameter (γ) is estimated while there are 4 when power is estimated separately on ages 6-9. In all cases AIC favours the traditional method (**sm*3p**). Equation 1 leads to the lowest biomass estimate of smh runs but the parameter γ is rather unstable in retrospective runs. The stock has to go through periods of large and small stock size (like a biomass-dynamic model) to estimate this kind of parameter while sufficient contrast is more easily obtained for age-disaggregated powers.

3.2.4.2 Overview

A summary overview of the models explored here (Figure 3.24, not including tunings with the fall survey only) show that none of them capture the rapid changes in observed data particularly well, although the overall trend is reflected by all models. The high observed spring survey biomass (smb) in 2012-2017 is well above predictions and the observed value in 2020 is between 72 and 85% of the predicted biomass. The fall survey measurements have shown a substantial variability in recent 5 years, with observed biomass in some years being well above or below that predicted.

A summary overview of the range of outcomes for the reference biomass (Figure 3.25), highlights the range of outcomes in any given past assessment year to the various setups explored. Tuning with the spring survey (smb) alone, all the setups were extremely consistent until 2011 (6% between highest and lowest estimate). but the difference has increased to 15% in 2020. When tuning with the fall survey (smh) alone the sensitivity of the reference biomass to the various setups explored has also increased since 2010 and is usually more than in the spring survey. When tuning with both surveys (smx) the same pattern of increased sensitivity to model setup with time is observed. Generally, tuning with both surveys show better congruence with the spring survey than with the fall survey. The difference between highest and lowest biomass estimate in 2020 is around 16% within each of the survey tuning scenarios (smb-range: 918-1061 kt, smx-range: 965-1123 kt, smh-range: 1087-1269 kt). In the 2010 and 2015 assessment years the difference between the highest and lowest runs tuned with smb is 6 and 7%, 11 and 20% with smh and 8 and 10% with smx.

The results in Figure 3.25 show clearly that the factors accounted for in the various setups did not have much effects on the assessment before 2014.

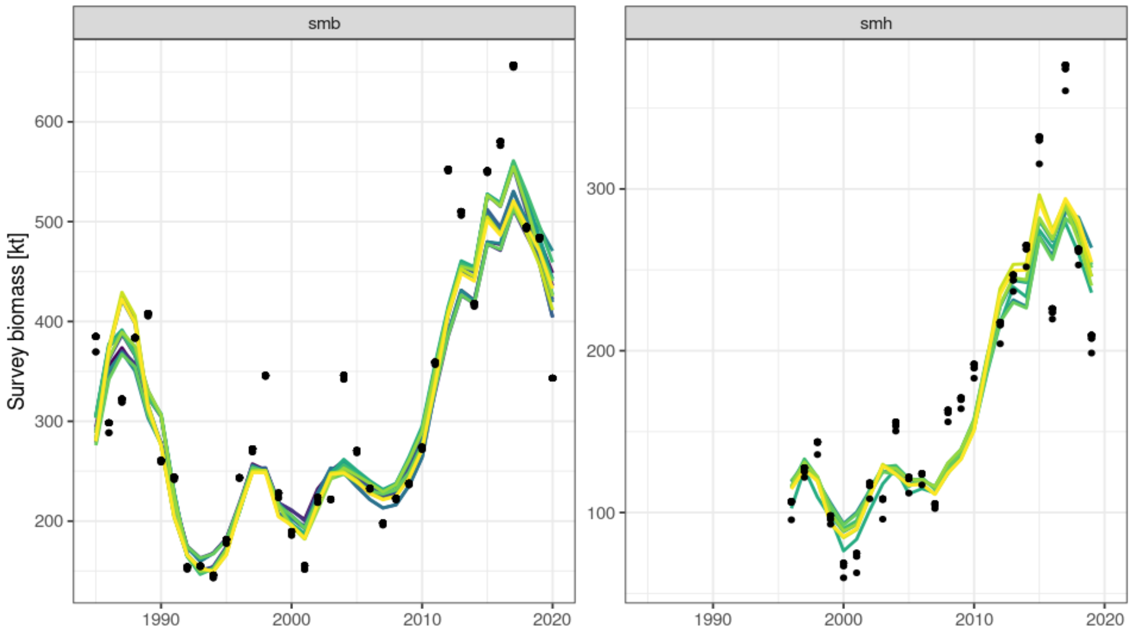

Figure 3.24: Observed (dots) and predicted biomass in the spring (smb) and fall survey (smh) for different runs (not showing results were the fall survey is only used in the tuning). The multiple dots within a year is a reflection of different age ranges being used in different runs.

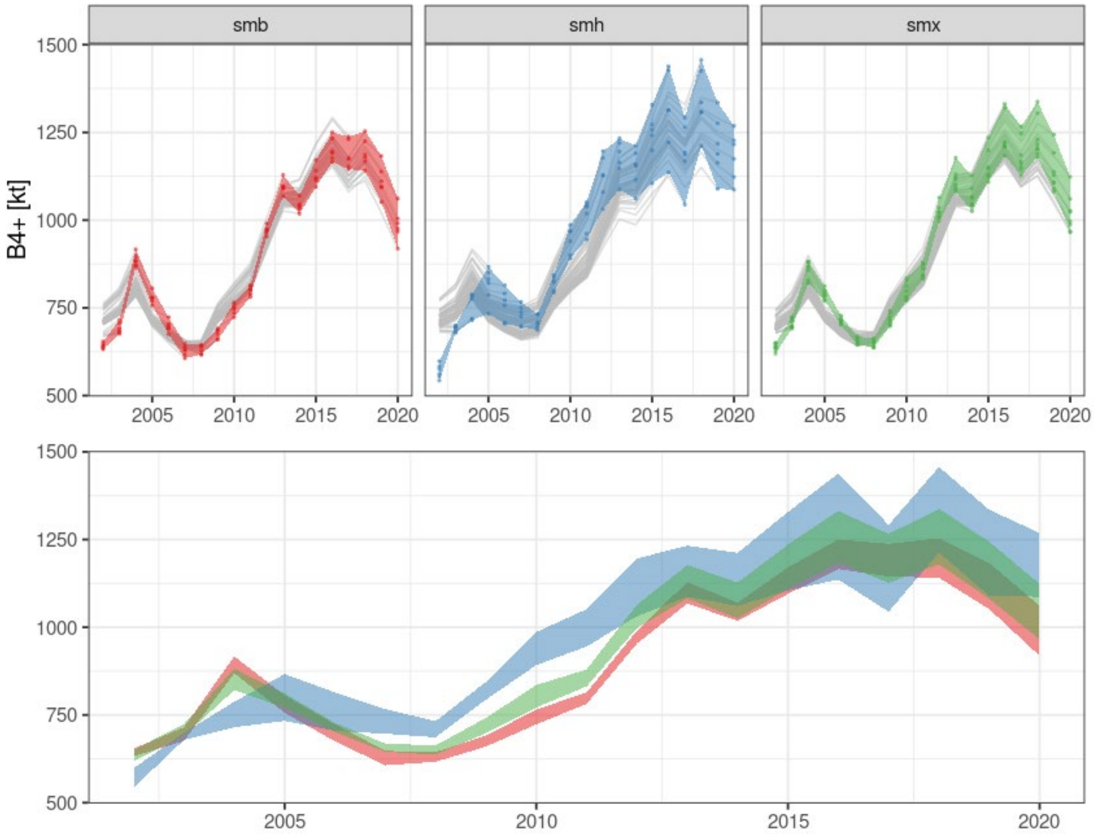

Figure 3.25: Estimates of reference biomass (B4+) in different assessment years based on tuning with the spring survey only (smb), fall survey only (smh) and both surveys (smx) for all the model configurations explored. Upper panels show the tuning survey combinations separately with the conventional retrospectives (dots the terminal value, grey lines the individual peels), the ribbon representing the range within the terminal assessment years. In the lower panel the ranges are superimposed for each of the survey combinations.

An overview of key statistics and outcome of the various runs in provided in the Table 3.1:

Table 3.1: Bias, autocorrelation and CV of reference biomass in retrospective runs from the assessments 2001-2015. Earlier assessments cannot be used as they have not converged. Assessments based only on SMB can be extended further back, Estimates of B4+, and size of year classes 2017 and 2018 from the 2020 assessment are also shown. Number of parameters is set to 0 for the lowest number of parameters where objective function is comparable followed by 1 or 2 runs with higher number of parameters and lower objective function

	run	bias	acf	cv	B4+	YC2017	YC2018	Objective	npar
1	smb1	0.016	0.266	0.030	1061	127	193	-1638.5	0
2	smb1p	0.005	0.316	0.027	977	125	187	-1660.5	4
3	smb3	-0.004	0.508	0.041	1004	129	129	-1593.0	0
4	smb3p	-0.011	0.540	0.037	918	128	128	-1616.0	4
14	smb3pa	-0.002	0.356	0.031	967	128	128	-1598.9	1
9	smx1	0.049	0.458	0.035	1123	139	204	-2010.4	0
10	smx1p	0.031	0.448	0.030	1026	136	196	-2046.8	8
11	smx3	0.028	0.569	0.040	1060	131	130	-1916.3	0
12	smx3p	0.016	0.503	0.033	968	130	130	-1954.0	8
15	smx3pa	0.023	0.511	0.033	988	130	129	-1930.6	2
5	smh1	0.106	0.582	0.064	1269	161	135	-1476.7	0
6	smh1p	0.098	0.692	0.073	1174	158	134	-1495.3	5
7	smh3	0.086	0.443	0.058	1216	133	133	-1425.3	0
8	smh3p	0.085	0.606	0.071	1123	132	132	-1442.7	5
13	smh3pa	0.012	0.697	0.063	1087	131	131	-1438.0	2
16	smx3parw12	0.032	0.483	0.031	1023	130	153	-1945.8	0
17	smb3parw12	0.003	0.263	0.028	992	123	149	-1608.3	0
18	smh3parw12	0.040	0.730	0.059	1227	152	135	-1445.2	0
19	smx1_3p	0.020	0.419	0.028	996	128	189	-1998.3	0

3.2.5 Extending the power relationship

In past assessment setups a non-linear relationship between survey indices and stock in numbers is assumed for age groups 1 to 5 in the spring survey but age 1 to 4 in the fall survey. Linearity is assumed in older age groups. This setup has been in place since the assessments in the early 1990's. The relationship between younger fish in the surveys and final number in a cohort has always indicated that not all fish were in the survey area and that log-log relationship could be replaced by line on normal scale with intercept (Figure 3.26). The value of q is also rather low leading to the same conclusions. Catchability of age 2 cod is less than 15% of that for adult cod and the number of age 2 cod is probably underestimated (true M much higher than 0.2). The same could be argued for age 1 which has even lower catchability, but their mesh penetration in the survey trawl could be a larger problem.

As expressed earlier, in the last 10 years there has been an increasing abundance of fishes older than age 5 (see Figure 3.7) resulting in an expansion in the ranges of observed survey indices and hence estimated stock in numbers beyond that previously observed (Figure 3.26). Although the observations are yet few and are mostly within the non-converged part of the assessment, there are indications that the non-linear assumption in older age groups may not hold true.

Coastal waters around Iceland have often been considered the nursery ground for cod (like for saithe) and one thing that could explain non-linearity in younger fish is that for larger cohorts, space will be limited closer to shore. Discovery of what explains the nonlinearity of older fish remains, however, illusive. The main problem with age dis-aggregated indices of older fish is to define the physical relationship behind a power relationship. To answer this question, we might have to go to individual stations in the surveys, the relationship between the numbers caught and amount of fish is most likely nonlinear, with higher proportions caught from dense schools. In the period of highest survey biomass, the calculated CV in smb was low or close to 0.1 except in 2012 when it was 0.22 (2012 was also a huge increase from 2011 - see Figure 3.6). A survey index is a sum or mean of the catches over the station and non-linearity at station level is lost when compiling the index.

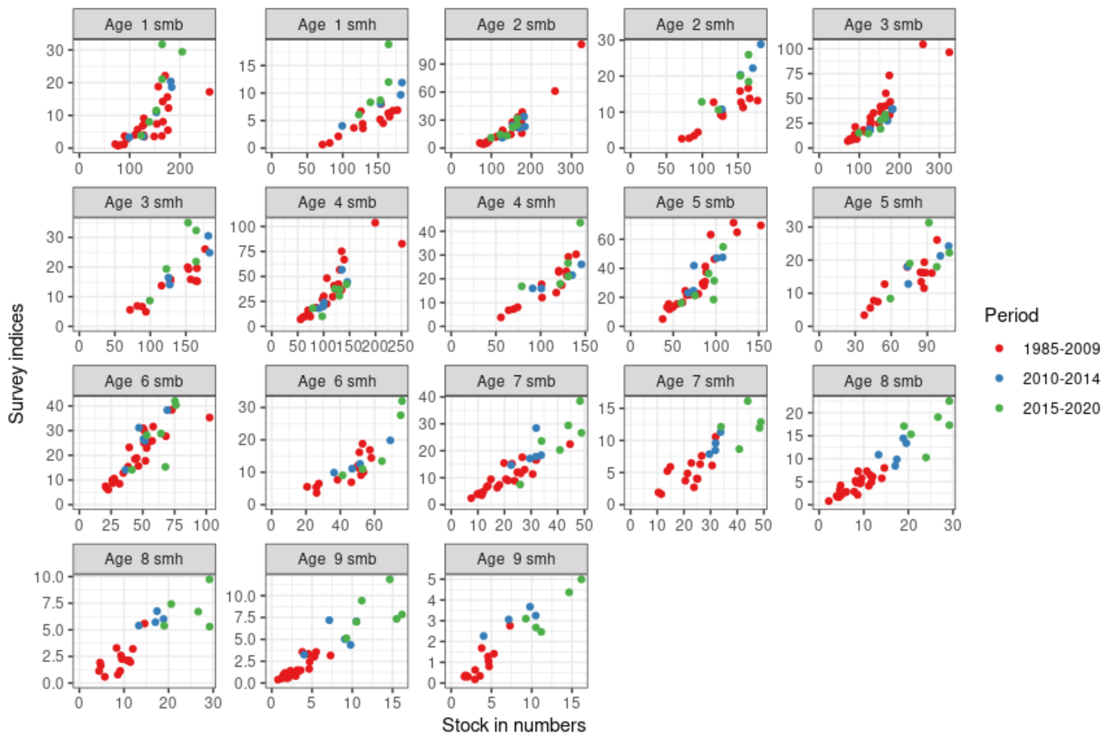

Figure 3.26: Stock in numbers and spring (smb) and fall (smh) survey indices by age (numerical value in each panel) based on the 2020 assessment setup (smx1) using muppet. The periods highlight the expansion in the range of survey and stock estimates in the last 10 years.

Based on tuning with both surveys, extending the power range to include ages up to 9 decreases the objective function from -2010.4 (**smx1**) to -2046.8 (**smx1p**) for 7 extra parameters, which is highly significant. This change also leads to a change in the estimated biomass in 2020, changing from 1123 to 1026 kt with improved retrospective patterns (Figure 3.27). Estimating power on total numbers (equation 1) leads to a higher AIC compared to adding power on each age group. Power was not explored for ages 10+ based on results in Figure 3.16.

The residual patterns (Figure 3.28) show fewer discrepancies of the important age groups 6 to 10 over the period 2012 to 2018 in the spring survey (smb) when the power is extended to age 9 but it cannot compensate for the low 2020 survey, resulting in a large negative block in all age groups up to age 9 (none of the settings tried in the exploration does that anyhow). In the fall survey an obvious block of positive residuals is observed in the three youngest age groups, residuals being largely negative in the beginning of time series but positive in more recent years.

It should be noted here that the retrospective performance of the base muppet model (**smx1**) has improved compared with the adcam run using the same survey setup (Figure 3.27 vs 3.17), the bias metric being 4% compared to 9%. This is most likely attributed to the changes in the cv profile and the correlation model for survey residuals (see earlier section) rather than changing from the adcam module to the muppet module.

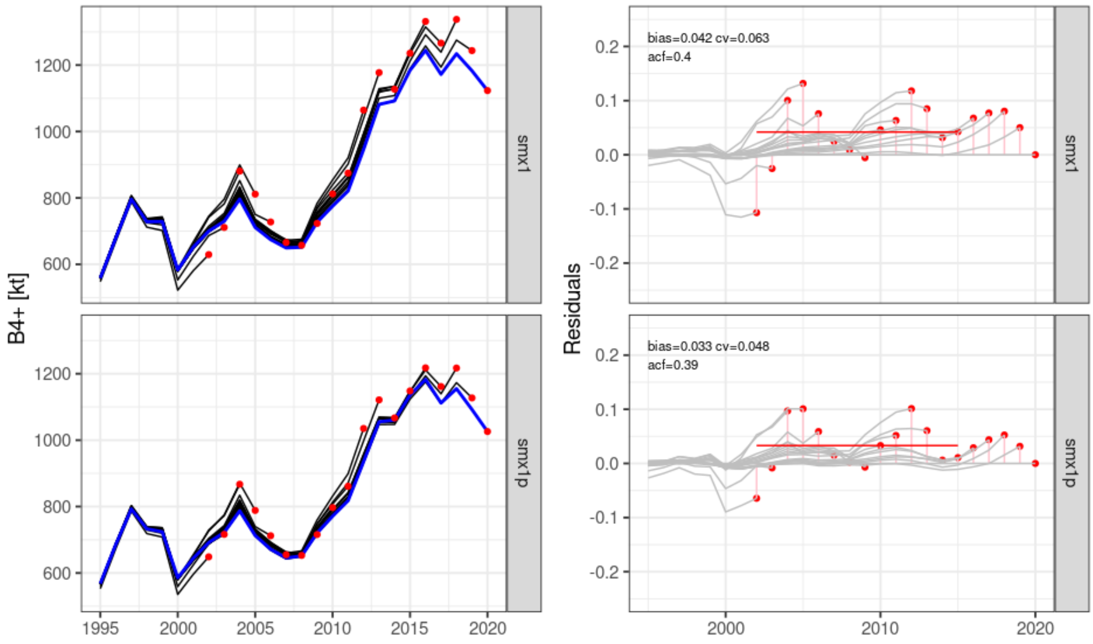

Figure 3.27: Estimates of reference biomass and retrospective patterns when tuned with both surveys ages 1 to 14. smx1: Power assumption up to age 5, smx1p: Power assumption up to age 9.

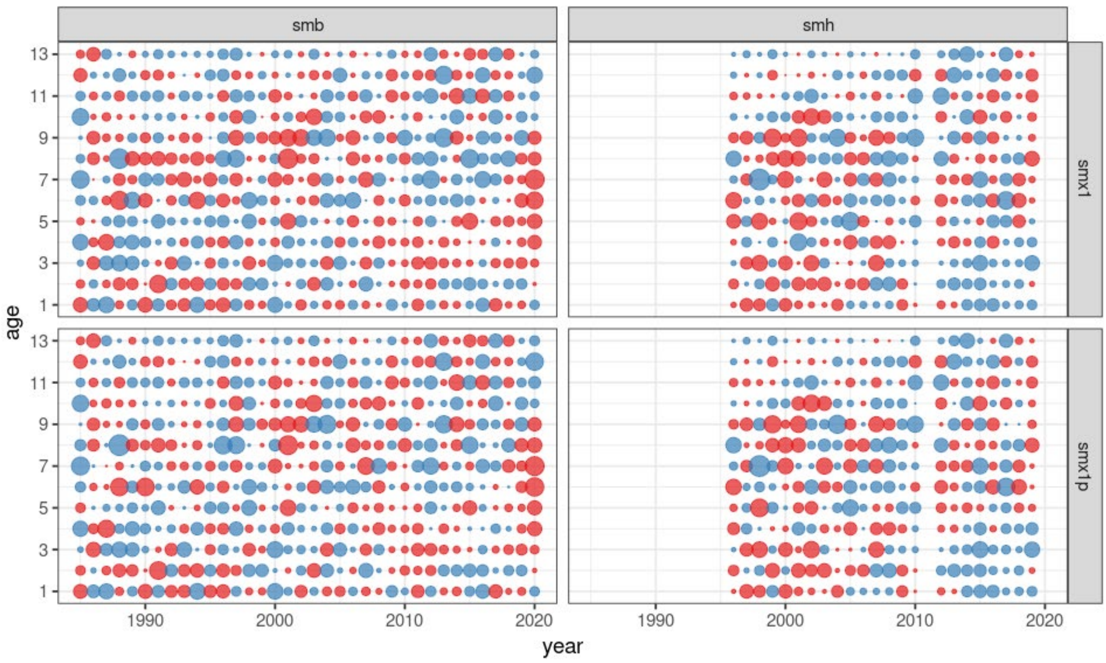

Figure 3.28: Residual patterns when tuned with both surveys ages 1 to 14. smx1: Power assumption up to age 5, smx1p: Power assumption up to age 9.

3.2.6 Dropping ages 1 and 2 from the fall survey

As noted above, an obvious block of residuals is observed in the three youngest age groups in the fall survey (Figure 3.28), residuals being negative in the beginning of time series but positive in more recent years, this being further illustrated in Figure 3.29. This led to exploring a run using both surveys, but where the two youngest ages from the fall survey were excluded from the tuning (referred to as **smx1_3p**).

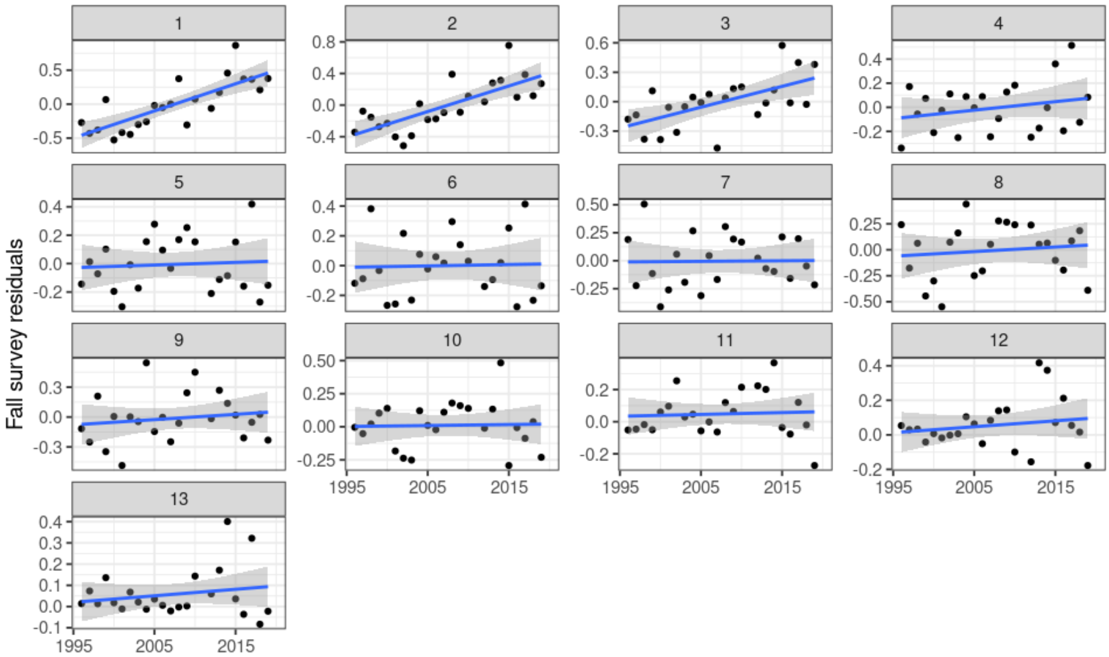

Figure 3.29: The fall (smh) survey residuals from the run using both surveys and power up to age 9 (run smx1p).

By dropping fall survey ages 1 and 2 from the tuning there is some improvement in the retrospective performance in the reference biomass estimate (Figure 3.30), both bias values (**smx1p** and **smx1_3p**) being within the acceptable range considered by ICES (more than 20% bias results in the assessment being questioned for being basis of advice). There are only minor changes to the terminal biomass estimate (1026 kt to 996 kt) implying that the influence of fall survey ages 1 and 2 on the reference biomass estimates is generally low. Retrospective patterns of other key dynamic metrics (Figure 3.31) as well as residual patterns (Figure 3.32) appear generally reasonable, although some time period blocks are still apparent in the fall survey for ages 3 and 4.

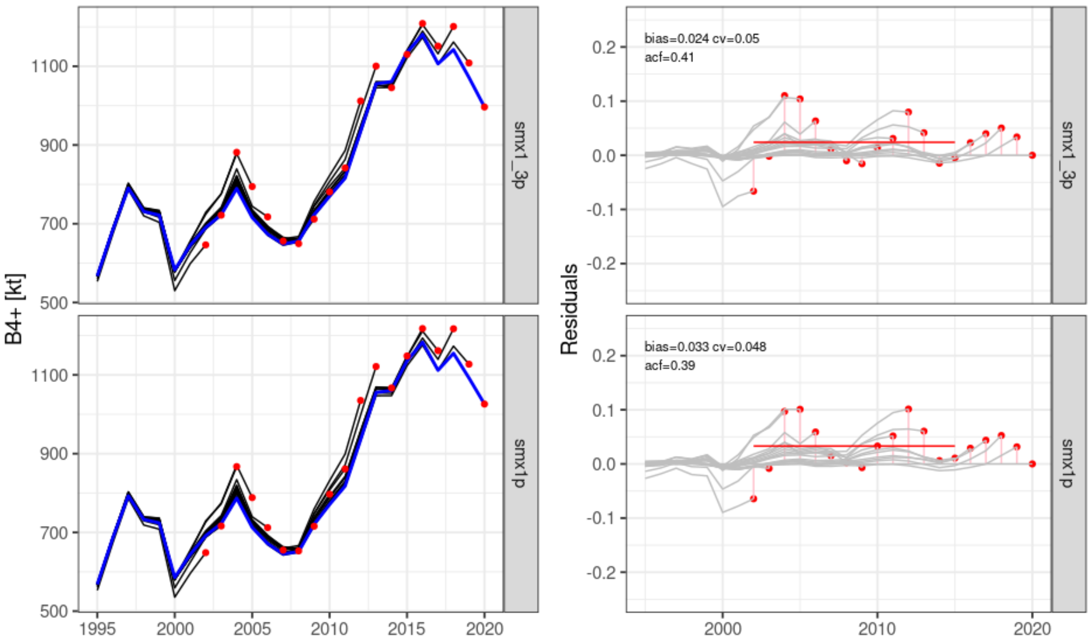

Figure 3.30: Retrospective patterns when tuned with both surveys and full age range with power assumption up to age 10 (smx1p, already shown in earlier figure) and same setup but excluding ages 1 and 2 from the fall survey (smx1_3p).

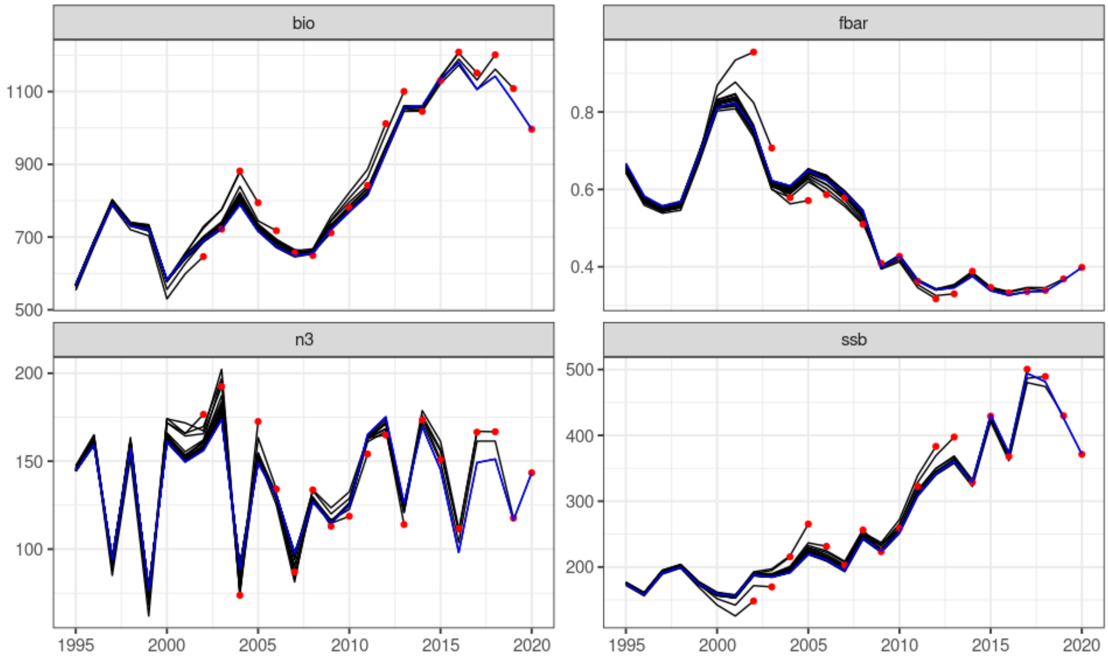

Figure 3.31: Results and retrospective patterns of key metrics when tuned with both surveys, spring survey ages 1 to 14, fall survey ages 3 to 13, with power assumption up to age 10 (smx1_3p).

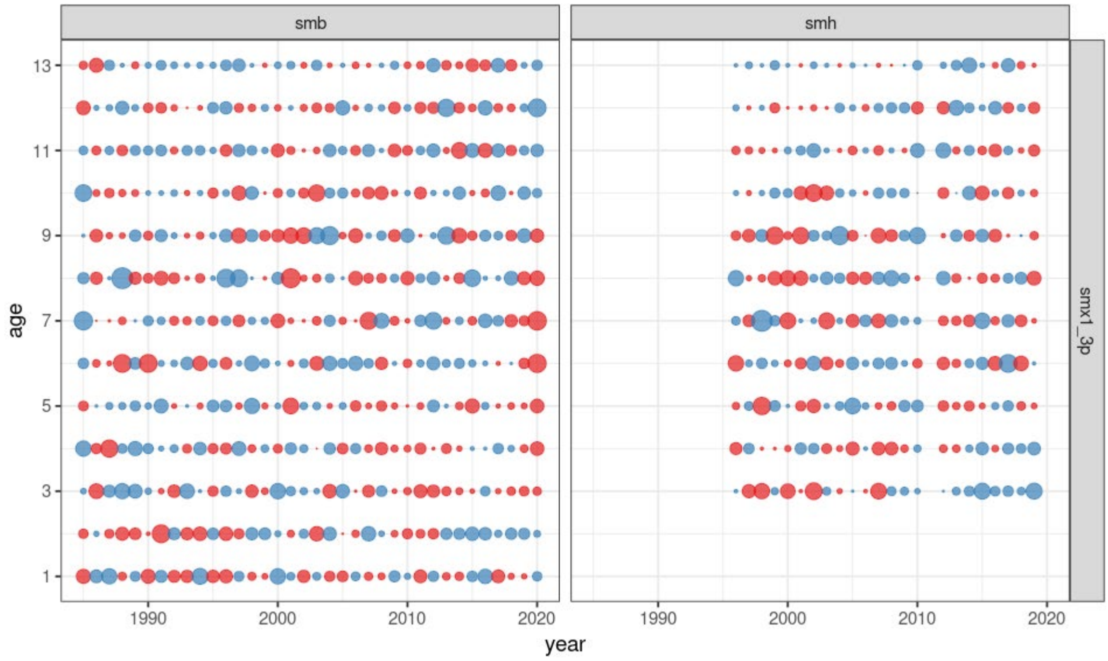

Figure 3.32: Residual patterns when tuned with both surveys, spring survey ages 1 to 14, fall survey ages 3 to 13, with power assumption up to age 9 (smx1_3p).

3.2.7 Should the assessment only be based on the spring survey or on both surveys?

As already noted (Figure 3.18) the adcam runs in recent years tuned with both surveys has resulted in the biomass trend falling in the middle range of values estimated when tuned with each survey separately. A comparison of the muppet assessments (**run1p**) results (Figure 3.33) show that the difference between tuning with the two surveys separately still show a difference of some 20% in the terminal estimates of the reference biomass but that the tuning with both surveys falls much closer to the spring survey tuning than the fall survey tuning. The higher congruence is most likely a combination of the updating of the cv-profile, the correlation model and the setting of power to ages up to age 9 rather than selecting a different model structure (adcam to muppet).

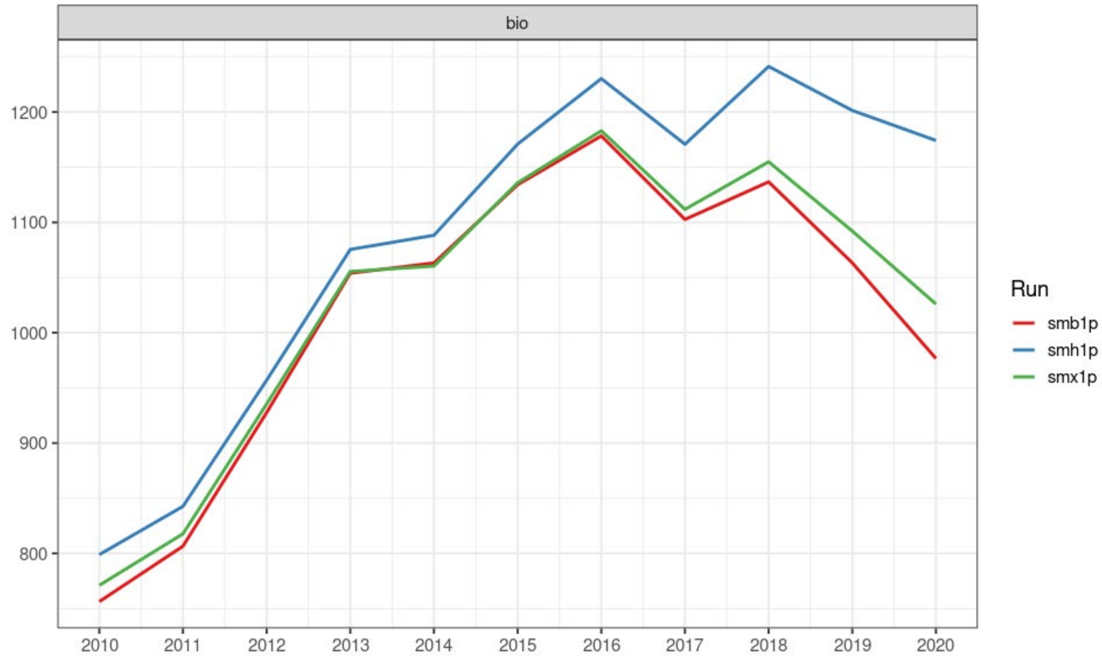

Figure 3.33: Comparison of the estimated reference biomass (bio) when tuning with the spring (smb) and fall (smh) surveys and when tuning with both surveys (smx). Based on the '1p' setting, i.e. where power is applied to ages 1 to 10.

Tuning with the spring survey alone (ages 1 to 14, power up to age 9, **smb1p**) results in less bias than when tuning with both surveys (ages 1 to 14 in spring survey, 3 to 13 in fall survey, power up to age 9, **smx1_3p**) (Figure 3.34). There are only changes to the terminal biomass estimate, 996 kt vs 977 kt compared with biomass value when tuning only with the fall survey (1174 kt, results not shown in figure).

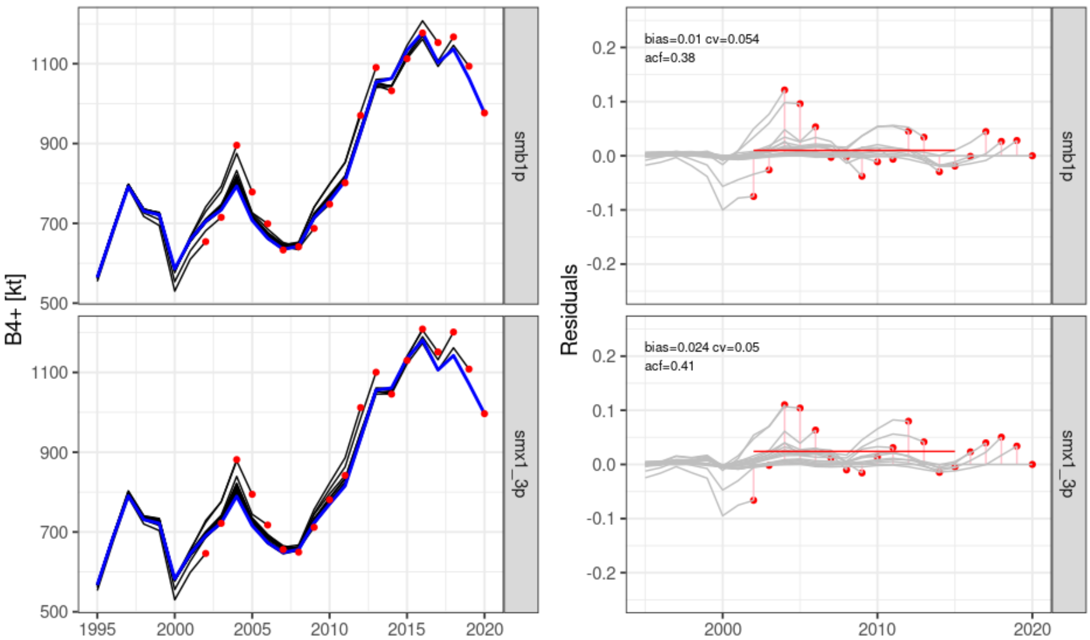

Figure 3.34: Retrospective patterns when tuned with both surveys (power to age 9, survey ages 1-14 for spring survey, ages 3-13 for fall survey (smx1_3p) and the spring survey alone (smb1p).

If an annual trend parameter is added to the module (using power on ages up to 9) the estimate is 0.4% per year for the spring survey but 1.5% for the fall survey. The latter estimate is significant but should not be taken too literally, such parameters should perhaps be estimated from VPA models and they might be confounded with assumed M. In most setups explored here, the retrospective patterns from the autumn survey show somewhat one-sided trend supporting the hypotheses that there could be a trend (results not shown).

Before the spring survey was started before 1985 the question was if the survey should be in March or October. One of the reasons that March was selected was that "vertical migrations" of cod were less in October (Pálsson 1989). Those vertical migrations could be linked to capelin predation that took place near the edges of the continental shelf north of Iceland. Looking at capelin from stomachs in the autumn survey, capelin amount and distribution show a decreasing trend possibly leading to more demersal behaviour of cod. Unfortunately, acoustic measurements have not been done in the autumn survey where commercial trawlers have often been used.

Another possible explanation is large negative year-effects in the fall 2000 and 2001 (same years in spring survey), close enough to the start of the fall survey to give trend. Assuming that there is not a trend but change in certain year it must happen between 2003 and 2007.

But trends in the survey can be related to using a incorrect observation model. The spring survey could also have trends in time although not yet detected. Candidates for causing trends are change in timing of cod spawning migrations, changes in spatial distribution of capelin in the spring, and even improved equipment to see that the trawl operates as it should.

3.2.8 Exploration based on dropping ages 1 and 2 from both surveys

The impact of excluding ages 1 and 2 from both surveys on the key stock metrics vs retaining them, (Figure 3.35) can be summarised as follows:

- The effect on the spawning stock biomass (ssb) and fishing mortality (fbar) is minor (although the scale range in the plots can be somewhat misleading).
- The main difference is in the recruitment estimates since 2014 (year classes 2011 and younger), being 9% lower when ages 1 and 2 are dropped from the surveys, this dissipating in an almost similar reduction in the reference biomass in 2020.

It should be noted that the absolute difference between the lowest and the highest 2020 biomass values considered in this section (108 kt) is well within assessment error used in the harvest control rule evaluation (see later).

Figure 3.35: Comparison of reference biomass (bio, kt), fishing mortality (fbar, mean ages 5-10), recruitment (r, millions scaled to age 3, x-axis refers to year class) and spawning stock biomass (ssb, kt) from runs where ages 1 and 2 are kept (sm1p) or dropped (sm3p) for tuning with the spring survey only (smb) or both surveys (smx). Run where age groups 1 and 2 were dropped from fall survey only is also shown (smx1_3p).

3.2.9 Explorations based on SAM

To get a somewhat independent view, the runs named smb1p, smb3p, smx1, smx1p, smx3, smx3p and smx1_3p were also run using the SAM model. The setup is of course not identical but the blocks and the power in relationship between survey indices and number in stock are identical as is the survey correlation of residuals. Treatment of catches is different, development of fishing mortality in SAM is multivariable normal random walk on $log(F_{ay})$. Also, SAM does not have any extra constraint on total catch, which is used in muppet. The setting of q, power and survey correlation are as similar as possible to the muppet settings.

The results in terms of reference biomass show good consistency between the two models (Figure 3.36). Historically the models converge to nearly the same value, the discrepancy before 1990 is likely due to that part of the 1984-year class in muppet runs is introduced as migration in 1990 and thus not in the books prior to that. In some years when there is a rapid decrease in fishing mortality, $F_{5_{10}}$ for example in 1995 is estimated as only 60% of 1993 value, so the multivariate random walk on fishing mortality will have some effect in those years. In recent years SAM predicts higher biomass as it has the tools ("process error") to reduce the stock more between 2019 and 2020. The reduction in $B4+$ between 2019 and 2020 is between 100 and 130 kt and 210 and 250 kt between 2018 and 2020. Comparable numbers from muppet are 70-80 kt between 2019 and 2020 and 110-190 kt between 2018 and 2020. This is possibly caused by process error that while it is relatively small is not negligible and can allow the stock to change faster in the SAM.

The SAM model follows observed catches very well. Regarding process error "logSdLogN" is estimated between -2.4 and -2.5, implying that deviations in $log(N_{a,y})$ have a standard deviation between 0.08 and 0.1. Putting separate estimates of "logSdLogN" for ages 1 and 2 was attempted to see if process error took care of variability of M in younger fish. The differences in values and likelihood were very small so only 2 parameters were estimated in each run. No catch constraint is applied in the assessment year, but the model does not try to increase F between 2019 and 2020 to follow the 2020 survey better.

But even though everything is similar in terms of main metrics the selection pattern of SAM and Muppet are quite different with the SAM model predicting dome-shaped fishing mortality in earlier years (Figure 3.37). In those years, the survey indices for older age groups are just noise. The dome-shaped selection pattern does not matter in this period as 11-14 years old fish are only a very small part of the stock.

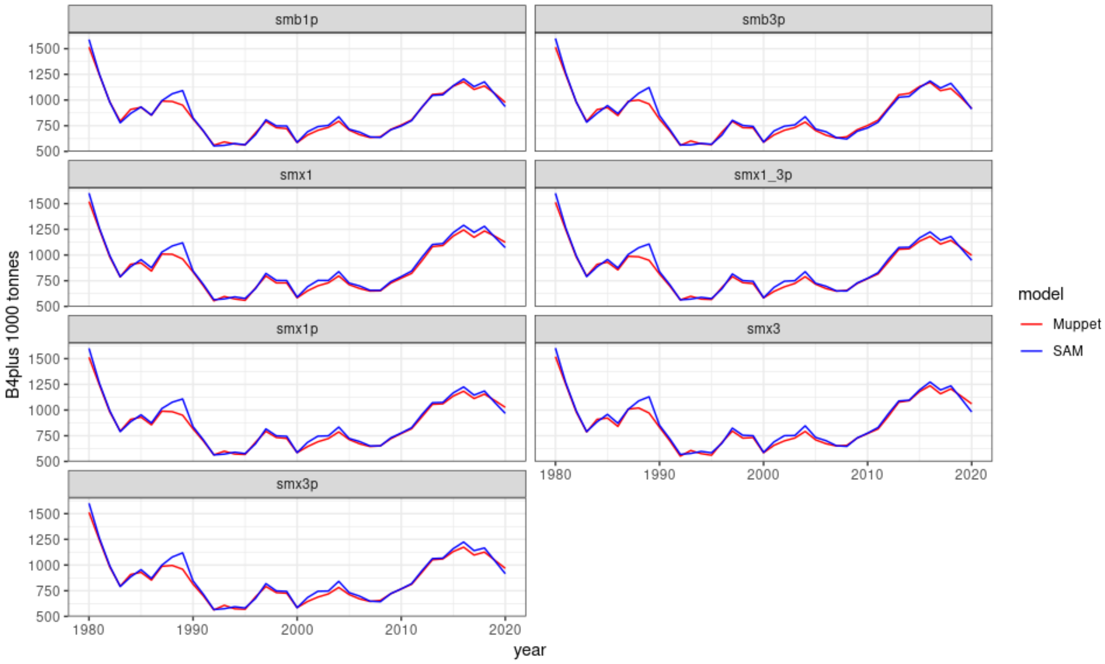

Figure 3.36: Comparison of muppet and sam using the same data input and relationship between stock-in-number and survey indices and survey correlation matrix. Panels refer to different setups, for explanation see text.

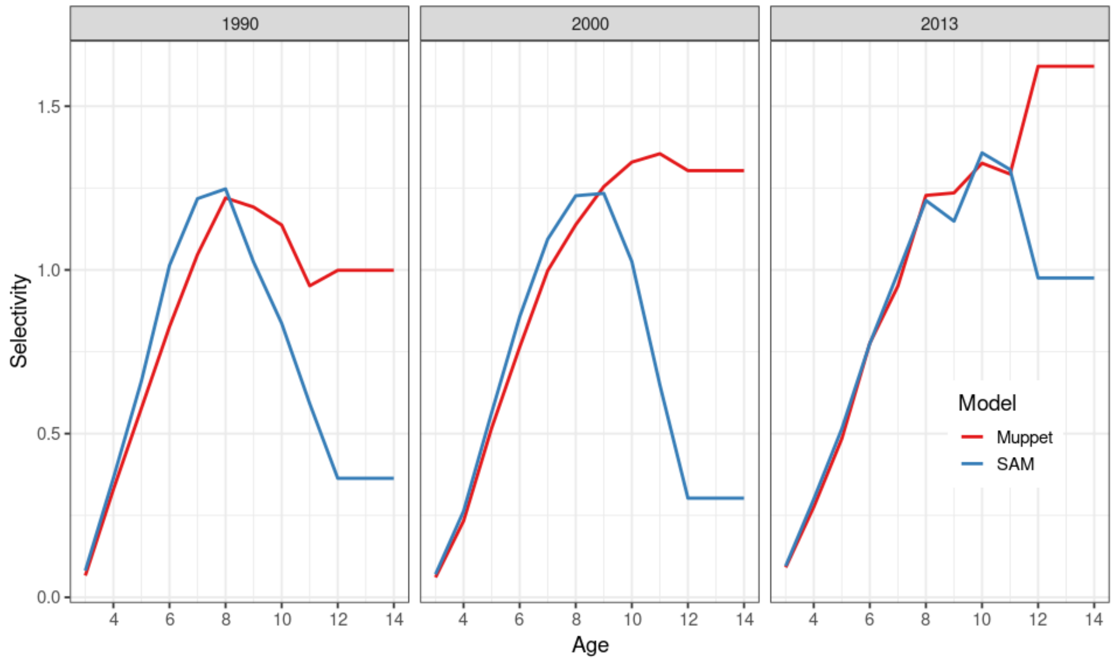

Figure 3.37: Comparison of selectivity (scaled to ages 5-10) from sam and muppet for selected years.

3.3 Summary

- The updating of the cv age profile and the autocorrelation and the inclusion of power up to age 9 resulted in the largest changes in the key metrics (recruitment, biomass and fbar estimates).
- These changes also meant that when using both surveys in the tuning, the effect of the fall survey was reduced such that now there is greater congruence of the combined tuning (smx) with the results than when only tuning with the spring survey (smb).
- With the same settings all models, adcam, muppet and SAM, give very similar results.
- A model setup tuning with spring survey ages 1 to 14 and fall survey ages 3 to 13 (**smx1_3p**) gave reasonable retrospective performance albeit a slight bias in reference biomass estimates (~2.4%). In terms of key stock metrics (biomass, recruitment and fishing mortality) this setup gave almost identical results as when tuning with the spring survey ages 1 to 14 only (**smb1p**).

4 Evaluation of potential harvest control rules

4.1 Summary of the simulation setup

Background

- **Motivation**: 5 year reevaluation of the HCR, check if current harvest rate is valid.
- **Main objectives**: In conformity with the precautionary approach and maximum sustainable yield.
- **Formal framework**: ICES on request from Iceland.
- **Who did the evaluation work**: WKICECOD 2021

Method

- **Software**: Assessment and simulation done within AD model builder. Downstream analysis and documentation done in R.
- **Name, brief outline**: Muppet. Statistical catch-at-age separable model, both the historical assessment and future simulations are done in the same framework. Assessment based on catch age 3-14 and two survey indices age 1-13 (smb) and 3-13 (smh). Shortcut approach (no assessment in the loop), historical retrospective errors applied to the true reference biomass (B_{4+}) and SSB when dictating future TAC.
- **Reference or documentation**: https://github.com/Hafro/Muppet_HCR
- **Type of stock**: Long lived, demersal, valuable
- **Knowledge base**: Analytical age based assessment.
- **Type of regulation**: TAC / ITQ.

Operating model conditioning

- **Recruitment**: Hockey-stick fitted to SR pairs 1955–2019. Lognormal distribution, SSB_{break}, CV and one productivity change (downward) estimated internally, autocorrelation applied externally based on estimates from model. Alternative hockey's explored.
- **Mean weight at age**: Average from 2011-2020, Deviations lognormal applied to all age groups within a year. CV and autocorrelation of the deviations based on average of estimated values for ages 4-10.
- **Maturity** Average over 2011-2020, no stochasticity. Alternatives explored, i.e stochasticity and correlated with mean weight at age.

- **Natural mortality**: Constant for all ages and year, no stochasticity.
- **Selectivity**: From the last separable period of the assessment (starting 2007), estimates from each historical iteration carried forward.
- **Initial stock numbers**: From the historical assessment, 2020 starting values from each iteration.
- **Decision basis**: Base case: TAC for next fishing year (starting September 1st in the assessment year) is a function of TAC in current fishing year (weight 0.5) and estimated B_{4+} in the beginning of the assessment year (weight 0.5). If SSB in the assessment year is estimated below management $B_{trigger}$, harvest rate is reduced and the stabiliser removed. The main characteristics is that TAC is only a function of B_{4+} in the assessment year so in principle no short-term prediction is required.

- **Number of iterations:** 2000.
- **Projection time:** 60 years.

Observation and implementation models

- **Assessment error**: Lognormal noise applied to the true reference biomass (B_{4+}) and SSB in the assessment year based on CV (0.12) and autocorrelation (0.52) estimated from the historical (empirical) assessment performance 1970-2015.
- Analytical retrospective patterns indicate lower CV than used here or ~0.05.
- Assessment bias explored.
- **Projection**: No projection needed to decide the TAC.
- **Implementation**: No implementation error assumed, effect evaluated qualitatively externally.

Harvest rule

- **Harvest rule design:**
- If $SSB_y < B_{trigger}$: $HR_{y/y+1} = HR_{std} \times SSB_y / SSB_{trigger}$
- If $SSB_y >= B_{trigger}$: $HR_{y/y+1} = HR_{std}$
- **Stabilizer:** If $SSB_y >= B_{trigger}$, final TAC is half the value of above and half of the TAC in the present fishing year. Not applied when $SSB_y < SSB_{trigger}$.

- **Duration of decisions:** Annual
- **Revision clause:** Normally evaluated every 5 years.

Presentation of results

- **Interest parameters**: Risk, average and stability of catch, ssb, reference biomass, recruitment and harvest rate (the mean and the 10th, 50th, and 90th percentiles).
- **Risk type and time interval**: Type 3, for decades 2020-2029 to 2070-2079. Short term considerations are most important here.

- **Precautionary risk level**: 5% of going below management $B_{trigger}$ and B_{lim}.

Experience and comments

- Variability in the historical assessment is not large, mostly related to variability in selection, migrations in some years (relatively uncertain). The variability in the initial stock size at the start of the simulation has some effect in the short term only, and the variability in the selection pattern carried forward is relatively minor relative to the assessment error applied.
- The evaluation is done relative to $B_{trigger}$, the B_{lim} is based on B_{loss}. The reason is related to productivity change ($\approx$ 35% decrease in recruitment) that seems to have occurred around 1985 SSB_{break} is estimated at B_{loss} (~120kt) that is used as B_{lim} both with productivity change included and not included. In earlier HCR evaluations if productivity change was not estimated, SSB_{break} was 2009 and before estimated at around 220 kt (that is current $B_{trigger}$), 171 kt in 2015 and at B_{loss} in the 2021 HCR evaluations. With productivity estimated SSB_{break} has always been estimated at B_{loss}.
  Comparing SSB_{break} at B_{loss} vs $B_{trigger}$ the values of the objective function are -1998.28 and -1991.71 with productivity change but -1986.916 and -1985.24 with no productivity change. Evidential support for the productivity change is therefore strong, the objective function changes by -12 for one estimated parameter that is a large change for $SSB - Rec$ parameters. Not including productivity change does not support $SSB_{break} = B_{loss}$ as strongly as when productivity change is estimated. Autocorrelation is estimated very low (~0.05) when productivity change is estimated but 0.2/0.3 when it is not estimated. Productivity change and $SSB_{break} = B_{loss}$ does therefore seem like the most plausible settings but no productivity change and $SSB_{break} = 220$ kt was closer in terms of objective

function in evaluations 2009 and earlier (2015 was in between).

- Trying a Beverton and Holt function with productivity change led to the same result as the Hockey stick, the slope was as high as allowed. Ricker function gives worse fit to the data than the Hockeystick, difference in objective function around +2. The base configuration Productivity change SSB_{break} fixed at B_{loss} fits the data best, not the original reason for selecting Hockey stick.

4.2 Description of the HCR evaluation setup

4.2.1 Biological parameters

4.2.1.1 Weights

The historical catch weights at age (Figure 4.1, see also Figure 4.2) show that there is some cyclical pattern in the mean weight at age and that deviations from the mean are more or less synchronous across principal age groups within a year, where the deviations are calculated via:

$$d = log(W_{a,y}) - log(\overline{W}_a)$$

The CV ranges from 0.07 to 0.20 increasing with increasing age (most likely a function of sample size). For the most important age groups in the fisheries and the reference biomass (4 to 10), the CV is around 0.08. A first order AR model (AR1) gives autocorrelation coefficient ranging from ~0.5 to ~0.7 (except the oldest age group).

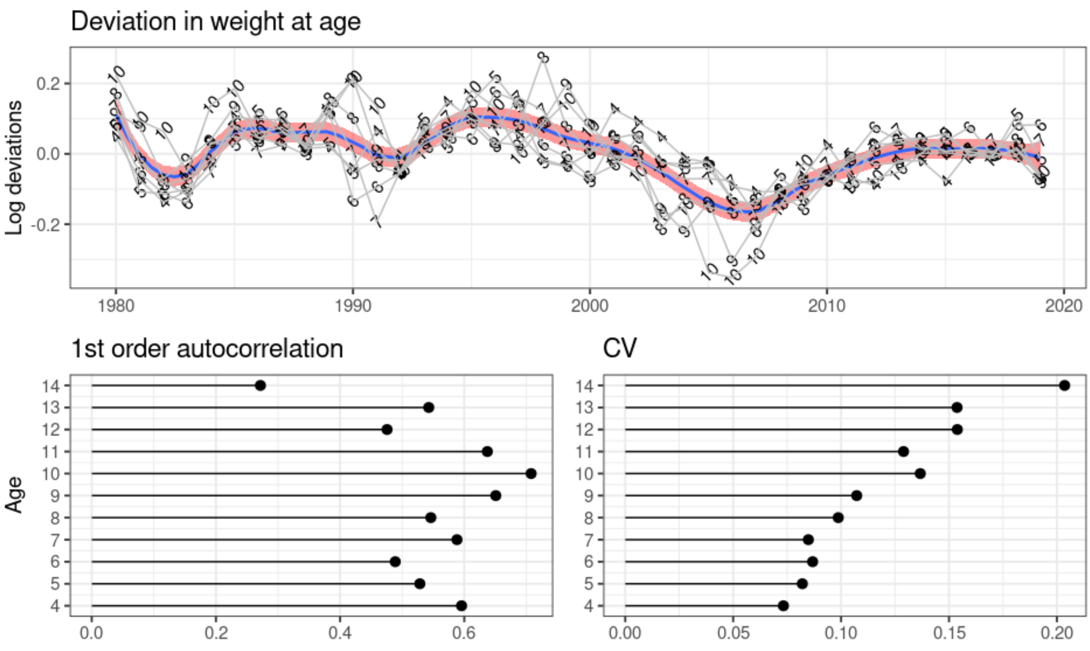

Figure 4.1: Catch weights at age 1980-2019. Deviation of log weight in each year from the mean log weight at age within each age group (age groups 4-10), Autocorrelation and CV of mean weight at age in the catches from 1989-2019.

The observation that principal age groups in the catches follow the same pattern in weight at age indicates that these age groups may experience similar food conditions although the actual driver is unknown. In Icelandic cod there are no or only very weak indication that weights at age are dependent on stock size, year class strength or the stock dynamics of capelin.

In the simulation it is assumed that the future mean weights are the average of those observed in years 2011-2020 with a $CV = 0.08$ and $\rho = 0.62$ applied across all age groups (Figure 4.2) within each iteration. The distribution of the CV and rho estimated within each iteration are shown in Figure 4.3. The basis $CV = 0.08$ and $\rho = 0.62$ was the average over agegroups 4-10 over the period since 1980. The weights 2011-2020 are close to the average 1980-2020.

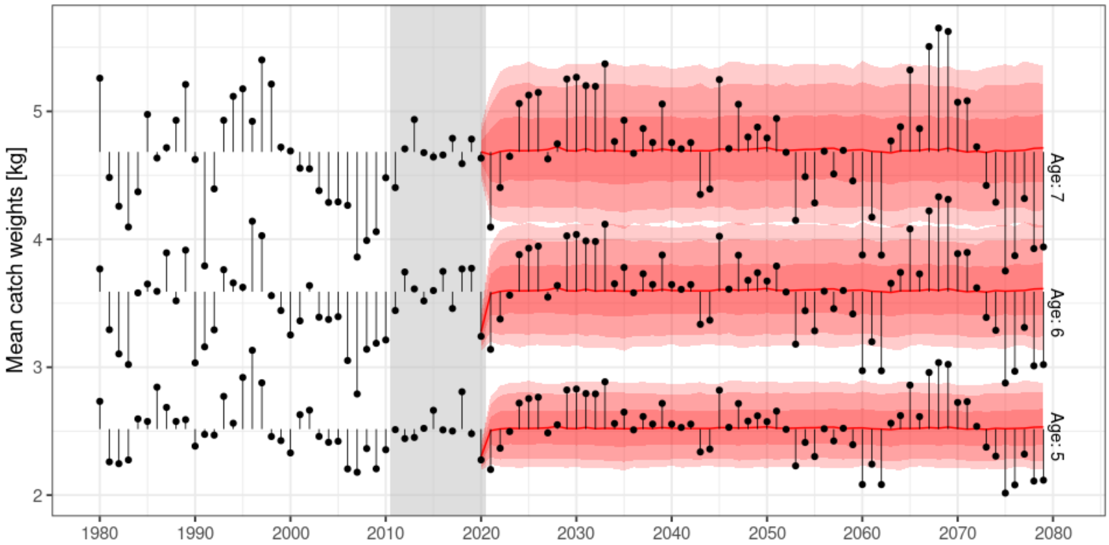

Figure 4.2: Historical weights at age of 5, 6, and 7-year-old fish in the catches in the last 40 years shown as deviation from the mean in the last 10 years (grey shaded area). Shaded red areas show the 50%, 80% and 90% distribution of values in the simulations and the red line the median. Also shown is one future iterations.

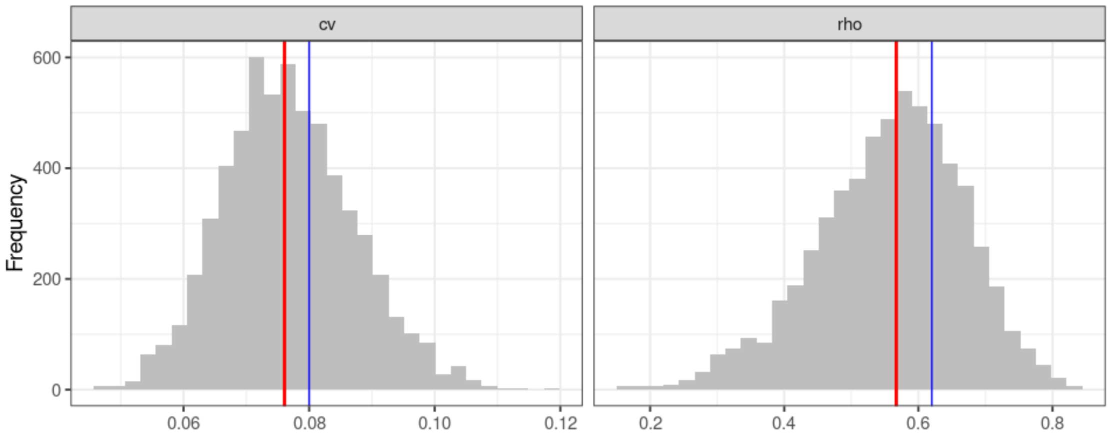

Figure 4.3: Distribution of CV (left panel) and rho (AR1, right panel) of the simulated weights with median values shown as vertical red line. Vertical blue line indicate the input values.

4.2.1.2 Maturity

The maturity at age in the spring groundfish survey is used as the basis to calculate the spawning stock biomass. After an increase in maturity at age in the first 15 years (1985-2000), there has been a decrease in maturity over the last 20 years in age groups 7 and younger while it has generally increased in the older fish over the whole period since 1985 (Figure 4.4).

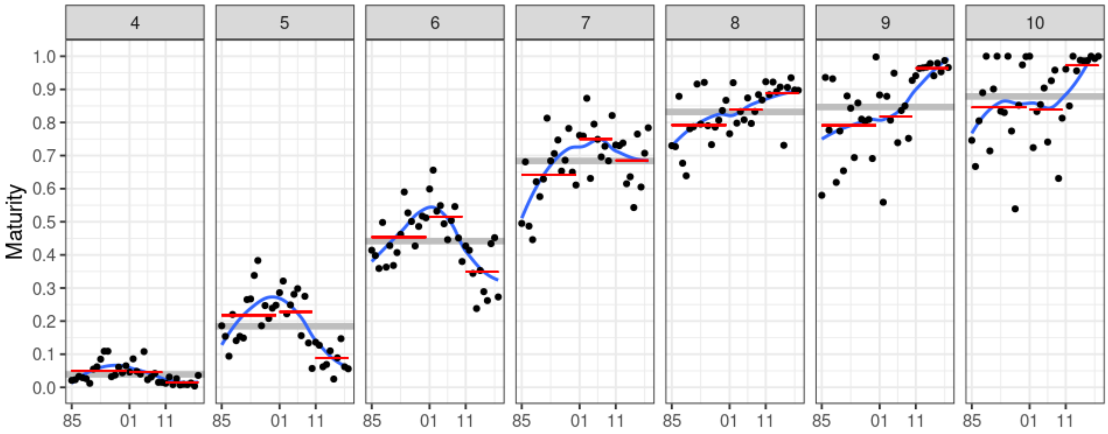

Figure 4.4: Maturity at age in 1985 to 2020. The red bars indicate the mean maturity over the time periods 1985-2000, 2001-2010 and 2011-2020, while the grey bar shows the mean over the period 1985-2020. The blue lines are a lowess smoother.

In the simulation the mean values were based on the last 10 years and in the base case kept fixed. A sensitivity analysis where the maturity was stochastic and correlated to mean weight at age. (Figure 4.5) was also run, effectively resulting in low maturity and weights coinciding in the simulation. Setting for this scenario were fitted to get something that "fitted reasonably" to data. The stochastic deviations in maturity were implemented as yearfactor, an incorrect but precautionary assumption. The maturity error is applied on logit scale with CV = 0.45 and autocorrelation = 0.7.

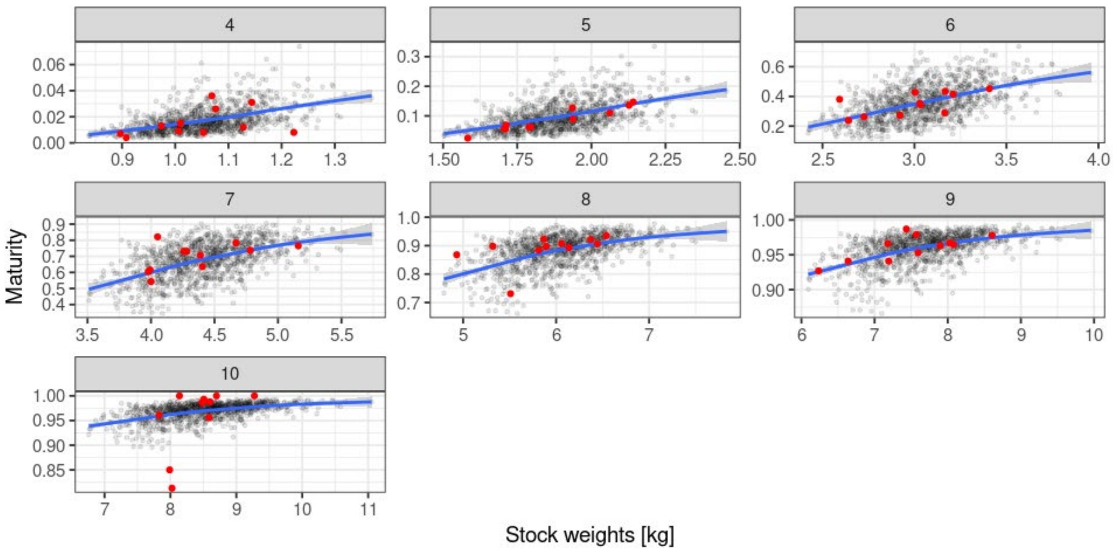

Figure 4.5: Stock weights and maturity for age groups 4 to 10. Red points are values from 2010 to 2020 and black points values from the simulation. A loess smoother if fitted to the simulated values.

4.2.1.3 Fishery selection at age

The selection pattern used in the simulation are the same as in the last separable period of the historical assessment (years 2007-2019) from each iteration (Figure 4.6). Effect of selection on key metric like stock biomass are relatively small as the decision rule used to set the TAC does not depend on selection in the analytical model but on the sum of product of the stock in numbers and catch weights of age 4 years and older (the reference biomass, B_{4+}).

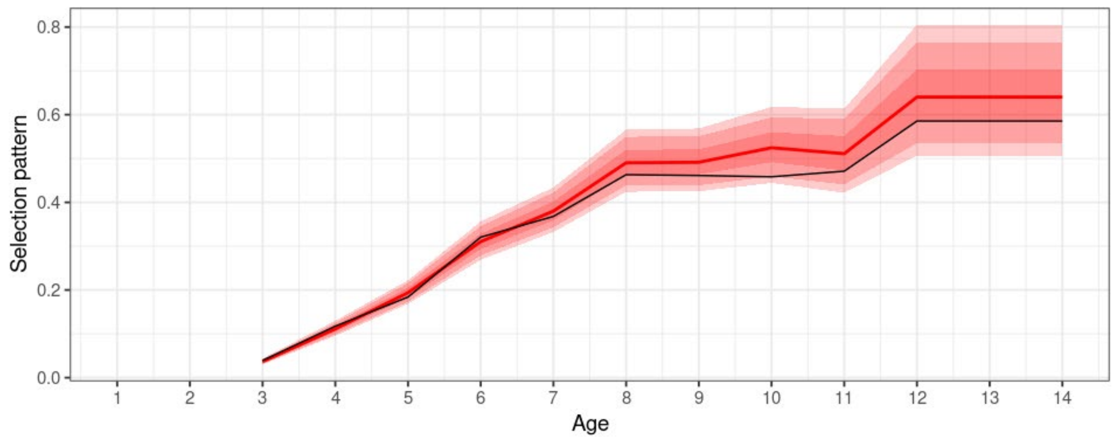

Figure 4.6: Distribution of the selection pattern used in the simulation. Shaded red areas show the 50%, 80% and 90% distribution of values in the simulations and the red line the median. Also shown is one future iterations.

4.2.1.4 Recruitment

The spawning stock size is estimated to have been quite large at the beginning of the time series in part driven by low fishing mortality during and after WWII (SCHOPKA (1994)) and a very large 1945 year-class at age 10 years and older in the beginning of the current time series. There is a decline in SSB until the middle of the 1970's when it reached a then historical low of 144 kt (year 1976). It generally remained below 220 kt until the end of the 2000's when it started to increase again, the current estimates indicating that it is similar in size as estimated in the mid 1960's (Figure 4.7).

Recruitment can be split into two distinct periods: The period 1955-1984 where the mean recruitment is around 200 million fish and the period after 1985 where the mean recruitment is only 130 million (35% lower). In the first period the recruitment in individual years is rarely below the mean recruitment of the latter period, while in the latter period recruitment in individual years never reaches the mean recruitment of the first period. The "break point" in the recruitment time series around 1985 does not correspond with a significant reduction in SSB, relative low stock size already observed in the early 1970's. No obvious oceanographic driver in Icelandic waters has been found that could be linked with the observed productivity change.

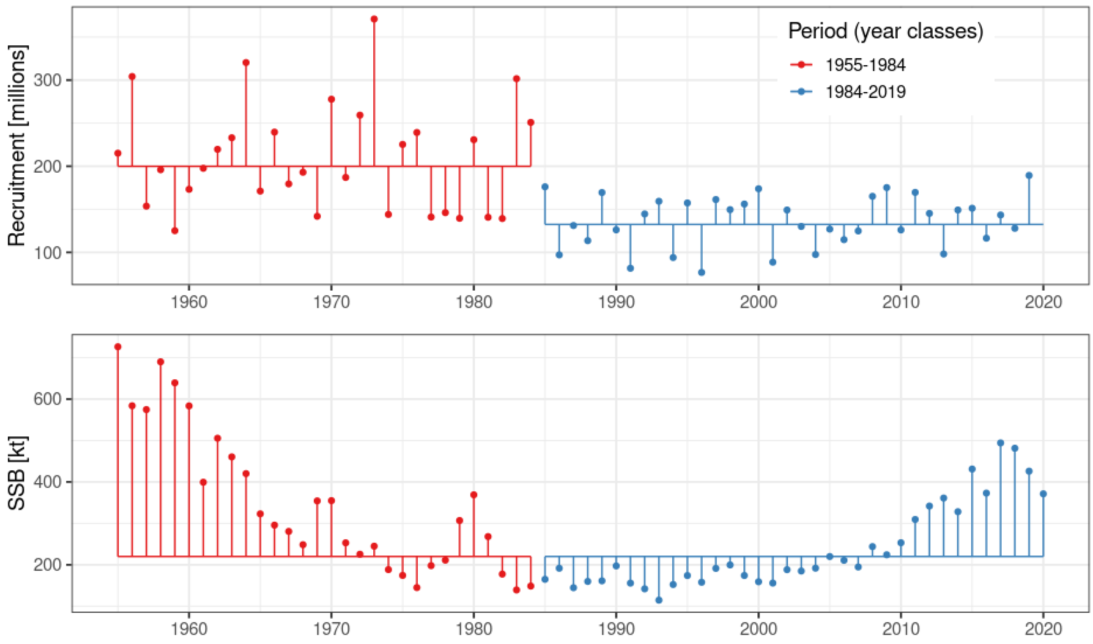

Figure 4.7: Recruitment (size at age 3) and spawning stock biomass over time. The recruitment before and after 1984-year class are shown as deviation from the mean within each period. The spawning stock biomass is shown as deviation from 220 kt. The x-axis on the recruitment plot refer to year class.

Year-class size as a function of spawning stock size, considering the full time series shows unsurprisingly, that variability is medium-high at any given size of the stock, and that large year-classes are still generated at the lower spawning stock sizes (Figure 4.8). There is some indication that the variability in recruitment increases with decreasing spawning stock size.

Taking the whole observation period there is an indication that the frequency of year-classes that are under the long term geometric mean (159 million) is somewhat higher when the stock is under ~220 kt (23 of 31 year-classes, or 74% of cases) than when the stock is over 220 kt (13 of 34 year classes, or 38% of cases). If one considers the patterns within two periods before and after 1985 there are no strong signals that the frequency of poor recruitment increases below the mean within each period when the spawning stock size is low.

Distinguishing between the hypotheses that the spawning stock biomass limits recruitment at low stock size, or if there has been a productivity regime shift around 1985, has been the main issue in all HCR evaluations since 2003. In the 2003 evaluations there were already indications of a productivity change around 1985 and the indications have become stronger with every new evaluation. Introducing a productivity change (multiplier on R_{max} after 1985) in the spawning stock recruitment function changes the objective (negative - loglikeli) function by -12 for one estimated parameter and gives a much better fit than can be obtained by any shape parameter (SSB_{break} for hockeystick , SSB_{max} or slope for Ricker). Significance level is often around -2 per additional parameter.

In the HCR evaluations up to and including that done in 2009, the SSB_{break} was estimated at around 200 kt if a productivity change was not included, in the 2015 evaluation it was estimated at 175 kt but is presently at or below B_{loss} (113 kt in the benchmark assessment). These changes in the estimate is a reflection of additional years of data being added that increases the statistical confidence of a permanent productivity change. Assuming productivity change has in all evaluation lead to $SSB_{break} \approx B_{loss}$. If a productivity change is assumed, autocorrelation of recruitment residuals is estimated to be small (~0.1) and CV is lower (~0.28) than if a productivity change is not included (CV ≈ 0.35, $\rho \approx 0.25$).

Comparing different HCR simulations conducted since 1994 the period after productivity change has gradually become a larger part of the total period

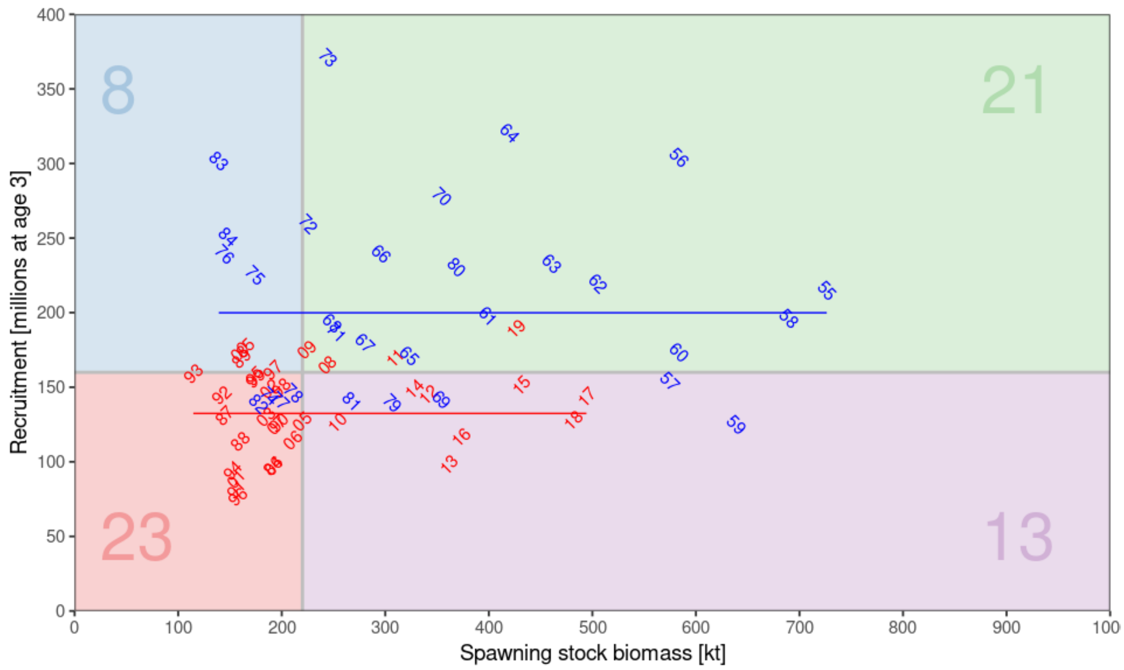

Figure 4.8: Recruitment as a function of spawning stock size. The numerical sloping labels refer to year classes, year classes 1955 to 1984 indicated in blue, year classes 1985 to 2018 in red. Horizontal lines show geometric mean for each period. Also shown are quarters split according to the long term mean recruitment (162 millions) and 200 thousand tonnes SSB, where the numbers indicate the number of year classes that fall within each quarter.

The HCR was tested using four variants of a Hockey-stick relationship (Figure 4.10):

1. Permanent change in the mean recruitment in 1985, estimated within the framework.

- Hockey 1 - the **base** case: In an initial evaluation, where the SSB-breakpoint was internally estimated in the framework, the observed values were well below any historical estimates. Setting a lower bound of 100 kt resulted all estimates being at or close to that bound. Hence the SSB-breakpoint was set externally at 120 kt (~Bloss). Autocorrelation in recruitment was estimated low (< 0.1) so it was fixed at 0.1 in the estimation.
- Hockey 1.1: A variant of the above, the SSB-breakpoint set at 200 kt. This is labeled as **b200** in subsequent figures.

2. Hockey 2: No change in productivity assumed. SSB-breakpoint and autocorrelation estimated internally within the model. This is labeled as **1p** in subsequent figures.

In the case where permanent change in the recruitment is estimated the data earlier than 1985 are included in the model estimates such as CV. In subsequent analyses within the presentation the earlier recruitment data are scaled to the recruitment level after 1985 based on the productivity change estimate in each iteration. The median productivity reduction is estimated to be ~0.34 when the SSB-break is set at 120 kt, but ~0.28 when set at 200 kt (Figure 4.9). Fixing the SSB-break at 200 kt results however in mean recruitment when the stock is above the SSB-breakpoint 10% higher than when the SSB-breakpoint is at B_{loss}. Fixing the SSB-break at 200 kt leads to larger value of the objective function (+6.5 worse fit) for same number of parameters, relatively large change for a parameter of a SSB-rec function. This is also manifested in that the historical probability profile is to the left of the simulated values.

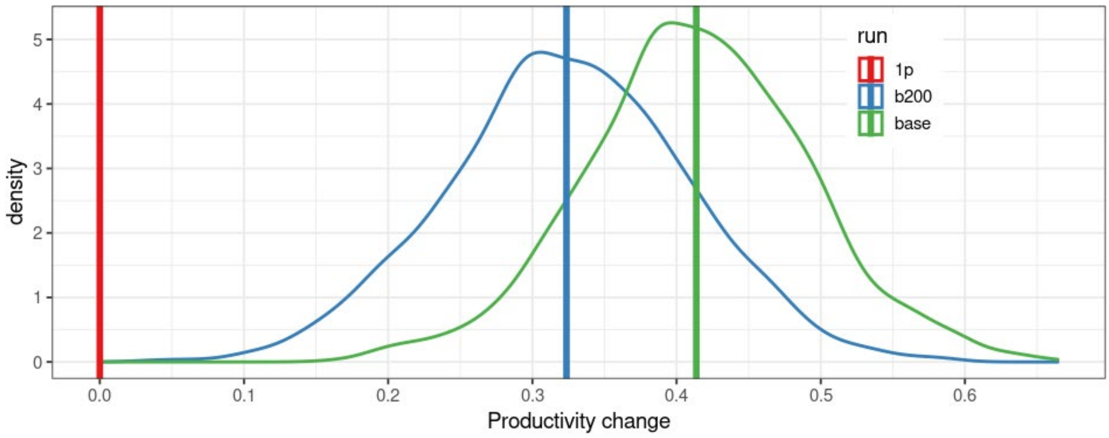

Figure 4.9: Estimated productivity change distribution and median from two hockey-stick recruitment scenarios. 'b200': Breakpoint fixed at 200 kt, 'base': Breakpoint lower bound set at 120 kt (essentially fixed).

The probability profile based on the different scenarios indicate (Figure 4.11, but see also Figure 4.10) a reasonable overlap in the historical and simulated recruitment distribution. Fixing the SSB-break at 200 (**b200**) case shows somewhat more variability in recruitment than the base case (**base**) as the stock more often enters the region of reduced recruitment and when the stock is large the recruitment is higher.

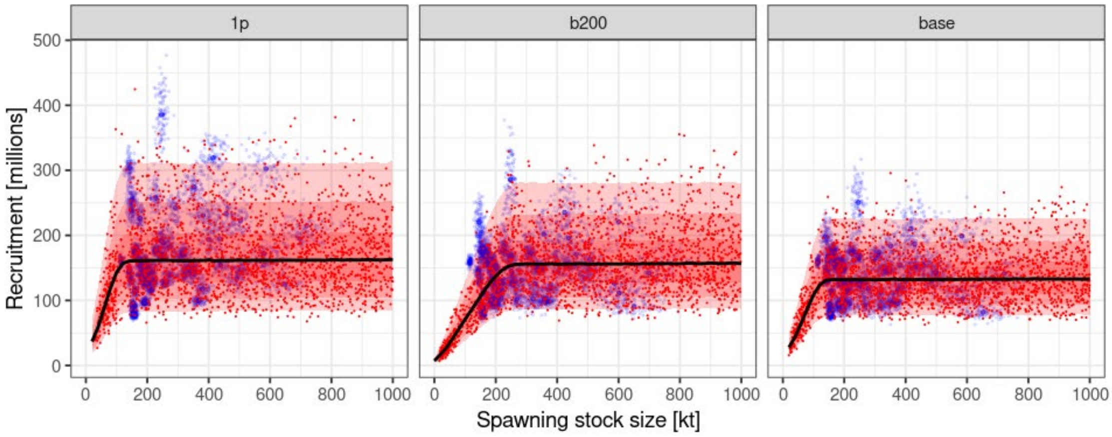

Figure 4.10: Recruitment as a function of spawning stock size. '1p': No productivity change, 'b200': Breakpoint fixed at 200 kt, 'base': Breakpoint lower bound set at 120 kt. Historical estimates (blue points, large points median value, small points 200 iteration samples. For simulations allowing productivity change break values prior to 1985 are scaled to the recent recruitments) and future samples (red points samples from iterations, ribbons showing 95%, 80% and 50% distribution and black line the median.

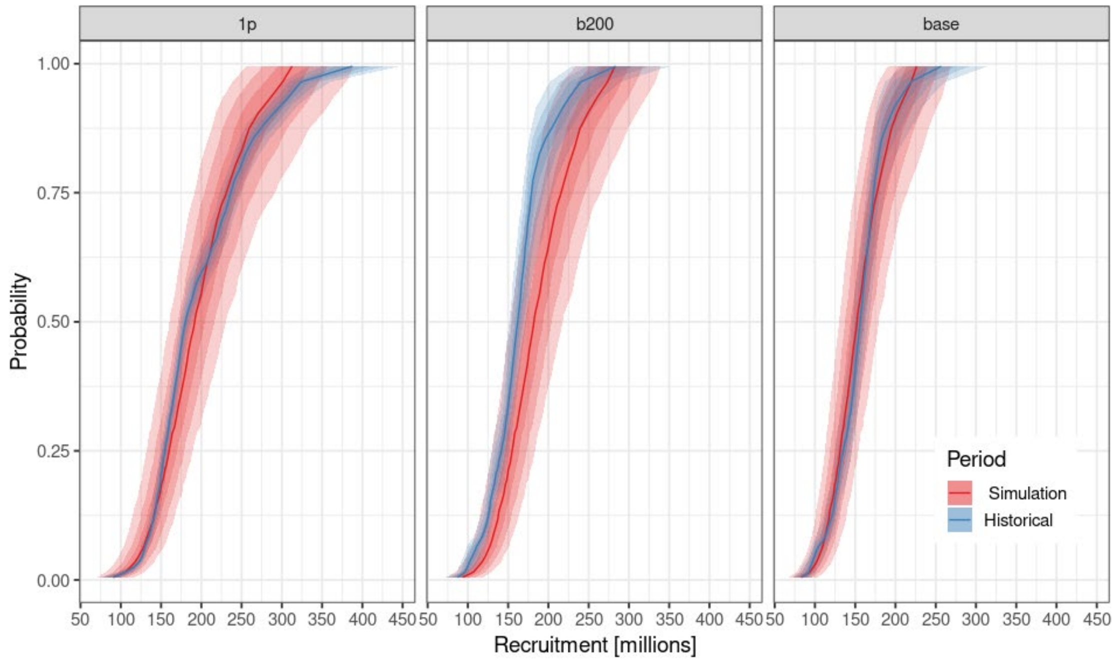

Figure 4.11: Probability profile of historical and simulated recruitment of the 3 different recruitment scenarios ('1p': No productivity change, 'b200': Breakpoint fixed at 200 kt, 'base': Breakpoint lower bound set at 120 kt) when fished in future at a target rate of 20%. Ribbons showing 95%, 80% and 50% distribution and lines the median value.

4.2.1.5 Natural mortality

Natural mortality is set to 0.2 in all age groups both in the assessment and the simulations. Natural mortality is without a doubt higher in the younger age groups than in at least the median age groups. The assumed value of M for ages 1 and 2 does not matter, since they are not caught and a value of 0 could just as well be used. Catch of age 3 is also relatively small so the value of M used in these younger age groups has minimal effect on the evaluation presented in this report. Estimating M in the current model setup led to a value of 0.21 (90% confidence intervals 0.18-0.23).

That said, it is recognized that incorrect assumptions about the scale of the natural mortality matters. That can though, only be explored when assessment is done "within-the-loop" something that is not done here.

4.2.2 Assessment error

Indication of assessment error, autocorrelation and/or bias can be obtained from three different sources:

- Comparison of the historical contemporaneous estimates irrespective data and model used with that of the current estimates (empirical retrospective).
- A retrospective evaluation using the current framework (analytical retrospective).
- Running an "assessment-in-the-loop" generating error in the survey and catch at age data. This method is not explored here.

The catch rule for the Icelandic cod dictates that the TAC in the advisory fishing year (starting 1. September in the assessment year (y) and ending 31. August in the next year (y+1)) is determined from the reference biomass ($B_{4+,y}$) and spawning stock in the start of the assessment year (y). The decision rule is thus not based on any prediction of the stock in numbers (beyond estimating catch-weights-at-age in year y from survey-weights-at-age in year y that are available). Hence estimates of assessment errors need only to be based on performance evaluation of the reference biomass in the start of the assessment year. Everything discussed here is based on the assumption that there exists something called a "converged assessment" i.e later reevaluations lead to more correct assessment.

4.2.2.1 Empirical history

The Icelandic cod has been assessed annually since 1970 (spanning now 50 years), in all cases using catch at age from age 3 to 14 and contemporaneous estimates of the terminal reference biomass (B_{4+}) is available since then. The source from the earlier years were obtained from various Icelandic assessment reports but since the early 1990's they have been reported annually in ICES NWWG reports. Initially the model used was some type of a VPA, often tuned with commercial effort series, followed by a combination of commercial and standardized survey indices in the 1990's. In the last 20 years the assessment has been based on a statistical catch at age model using only survey indices, first using the spring survey only but since 2010 the fall survey has also been included.

A comparison of contemporaneous assessments with the benchmark assessment (Figure 4.12) using values from 1970 to 2015 (converged part of the assessment) indicates that the CV is around 0.12 and the autocorrelation (AR1) around 0.54. The mean bias over the years 1970 to 2015 is around 7%.

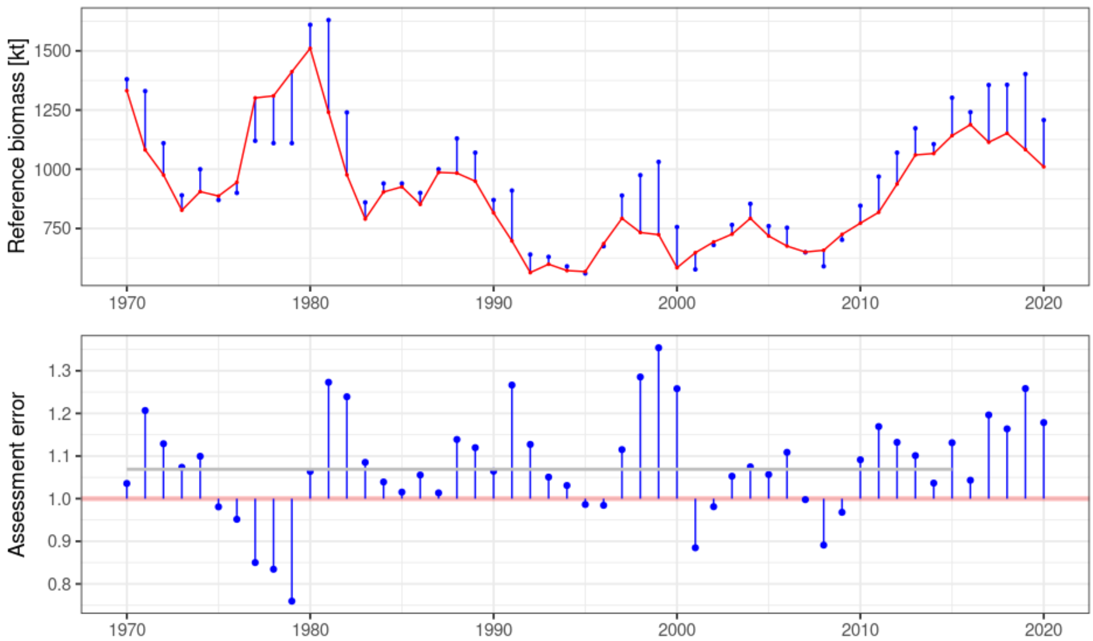

Figure 4.12: Historical assessment error. Upper panels are absolute estimates of the reference biomass (B4+) and lower panels relative to the terminal benchmark assessment. The red line is the benchmark reference biomass (B4+) and contemporaneous assessments are presented as blue points and lines. Horizonal grey line is the mean bias over the period 1970 to 2015.

4.2.2.2 Analytical retrospective patterns

The resulting CV in the analytical retrospective analysis (Figure 4.13) is much lower than that obtained empirically mostly because the current assessment setup has been made in response to problems with earlier assessments. There is an indication of a bias around 2.5% (which is well within the -0.2-0.2 criteria currently considered in ICES to be acceptable) but it is impossible to distinguish between that and potential assessment autocorrelation, given the inertia in the assessment that depreciates older data slowly.

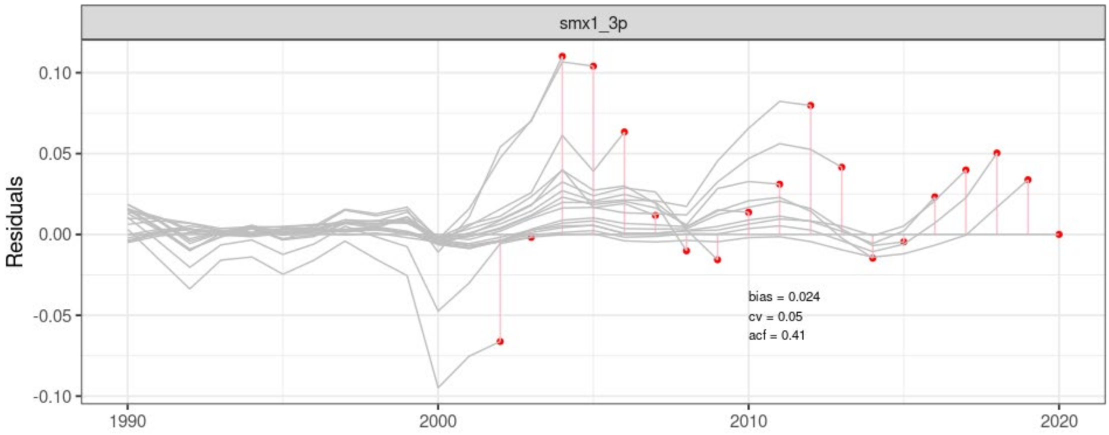

Figure 4.13: Analytical retrospective pattern of reference biomass of the 2021 benchmark setup.

4.2.2.3 Parameters used in the simulations

The simulations presented here were, as has been done since the 2003 HCR evaluation, based on the assessment error values obtained from the historical retrospective. This is something of a departure from the guidelines where the purist thinking is to derive the error structure from the current assessment framework (i.e. specific to a model and data). For the Icelandic cod we have the luxury that the assessment and hence advice, has since 1970, been based on a catch-at-age model using ages 3 to 14. Over time the difference in the assessments has thus been the model selected for use, the data source used for the tuning of the terminal estimates, and the fine tuning of the model setup (assumption on selection pattern, relationship assumed between tuning indices and stock in numbers and the relative weighting of the different inputs). Applying the estimates of cv and rho from the historical performance can hence be considered to include error associated with any future changes in the catch-at-age model tuning-setup resulting in the decision rule being more likely to hold over the long run.

The future distribution of the simulated assessment errors (Figure 4.14) shows that they broadly capture the observed historical range. Distribution of the estimates of the main metrics (Figure 4.15), calculated during each iteration indicate, that the median in the CV, auto-correlation and bias are approximately similar to the input values. The bias shown here is not the 7% bias in the historical retrospective analyses but the bias in the lognormal distribution ($e^{0.5\sigma^2}$).

In the simulation the error structure in the advice-model is applied as follows:

1. The "true" reference biomass ($B_{4+,y}$) is calculated from the stock-in-numbers at the beginning of the assessment year (y) and the estimated catch-weight at age in the assessment year (derived from prediction based on the spring survey weights in the assessment year).
2. The estimated reference biomass used in deciding the TAC is calculated as follows:

$\widehat{B_{4+,y}} = B_{4+,y}e^{\epsilon_y}$ where ϵ_y is the assessment error.

$$\widehat{SSB_y} = SSB_y e^{\epsilon_y}$$

3. The advised catch for the next fishing year ($Y_{y/y+1}$) is derived from:

If $\widehat{SSB_y} > B_{trigger}$

$$TAC_{y/y+1} = 0.5h\widehat{B_{4+,y}} + 0.5TAC_{y-1/y}$$

where h stands for the harvest rate and $TAC_{y-1/y}$ is the TAC from the previous year assessment.

If $\widehat{SSB_y} < B_{trigger}$

$$TAC_{y/y+1} = \frac{\widehat{SSB_y}}{B_{trigger}} h\widehat{B_{4+}}$$

4. The true stock in numbers are propagated through the assessment year using the F_{ay} derived from estimated total catch in the assessment year (2/3 of the TAC from fishing year $y-1/y$ and 1/3 from the fishing year $y/y+1$)) applying the selection pattern estimated in the first year of the simulation (year 2020 in this case).

A problem with this decision rule is that it responds very slowly when SSB is estimated above $B_{trigger}$, both because of the catch stabiliser and inertia in the assessment that depreciates old data slowly. Below $B_{trigger}$ target harvest rate is reduced, the stabiliser is turned off, and harvest rate has probably been higher when that happens so the model responds faster. This faster response reduces stability of catches and to avoid that, the HCR has internally been evaluated against $P(SSB < B_{trigger} = 220kt) < 0.05$ even though the formal risk evaluation is like for other stocks based on $P(SSB < B_{lim}) < 0.05$

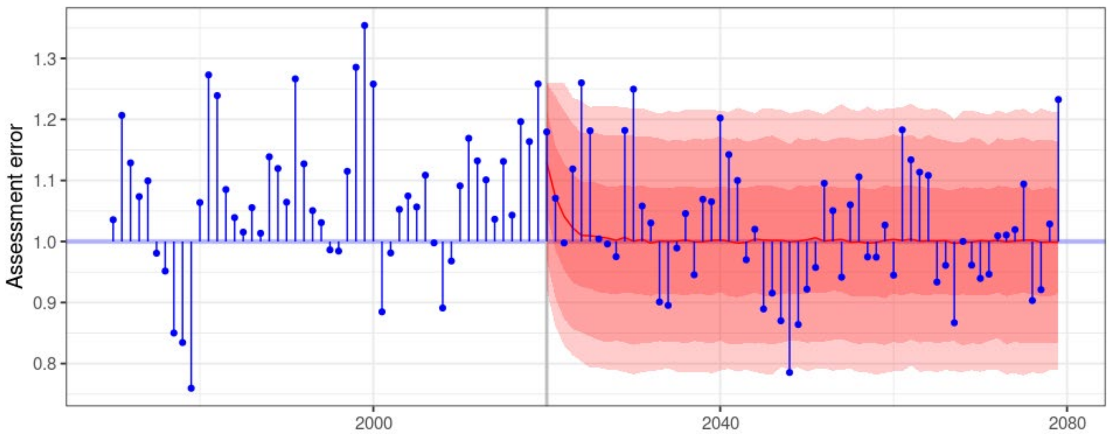

Figure 4.14: Historical assessment error (points and lines) and one future iterations (blue points and lines). Shaded red ribbons show 50%, 80% and 90% of the distribution.

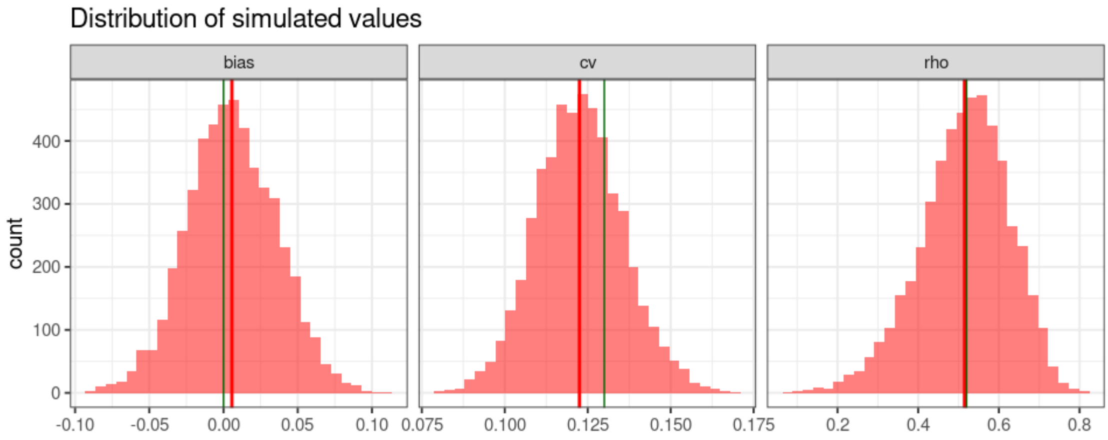

Figure 4.15: Distribution calculated from each iteration and the median (red vertical line) of the simulated metrics and the input values used (green vertical line).

4.2.3 Implementation error

A management decision on catch-constraint based on a formal harvest control rule was first made for fishing year 1995/96. In general, the landings have been above the catch advice (Figure 4.16), with the average bias being around 5.3%. The main reasons for the overshoot include:

- Catch of vessels less than 10 GRT that were not under the ITQ system were underestimated in the 1990's and the beginning of the 2000's. These are now also in the ITQ system.
- Account of catch of foreign vessels were improperly or not fully taken into account prior to allocating the remaining catch to the Icelandic ITQ vessels.
- A recent procedure to "call in" initial over-allocated catch to the ITQ system has not worked as well as intended.

Implementation bias is not taken into account in the future simulation but a 5% bias in catches at a 20% "target" harvest rate is equivalent to performance measures at harvest rate that is 1% higher (21% harvest rate).

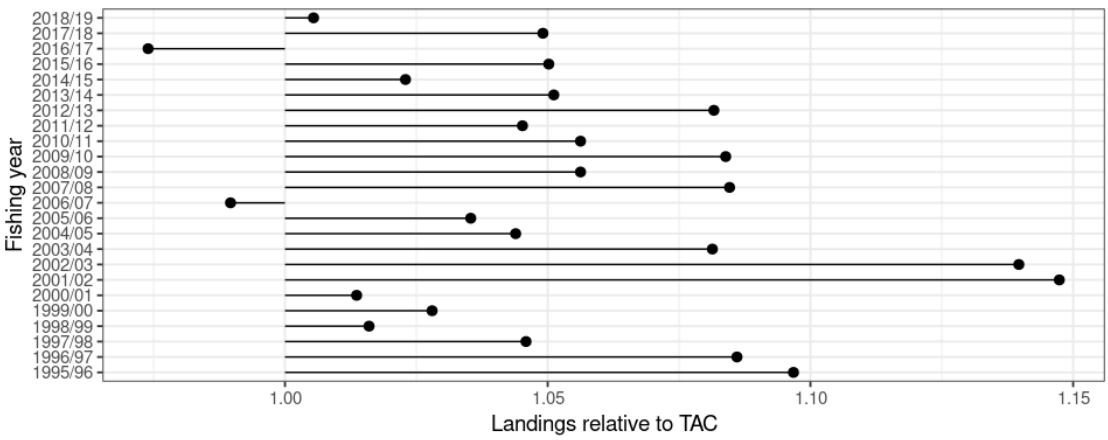

Figure 4.16: Landings relative to TAC.

4.3 Results

The general dynamic patterns from the simulation for difference target harvest rates (Figure 4.17) indicate, not un-expectedly, that a higher harvest rate leads to a smaller biomass while the general catch level remains similar but with higher variability. Variability in fishing pressure (realized harvest rate) increases also with higher target rate in the control rule, in part because recruitment variability has a greater influence on the reference biomass (is then composed of fewer and younger age-classes) and because the lag and the lead inertia in the catches relative to the biomass because of the buffer.

4.3.1 Long term equilibrium

The harvest rate leading to maximum catch is poorly defined (Figure 4.18), the point estimate of the target harvest rate that gives maximum yield for the base case being 0.23, but the maximum of the 5th percentile of catch being lower, around ~0.19. The profile of the spawning stock biomass indicates that there is less than a 5% probability that the spawning stock biomass goes below 220 kt in the long term at harvest rates less than or equal to 0.22. Those numbers are all compiled based on including the stabilizer and no trigger. Taking assessment bias of 7% (historical assessment bias) into account, the target harvest rate that gives the maximum yield is 0.22 and the 5% $B_{trigger}$ probability will be achieved at harvest rate less than or equal to ~0.21.

Alternative scenarios and robustness tests of the three recruitment scenarios and then stochastic maturity (mvar) and correlated maturity (mcor) based on the base recruitment scenario (Figure 4.19) show that the long-term maximum catch is obtained at a target harvest rate of ~0.22-0.23. The no-change-in-productivity scenario (1p) and when SSB_{break} is fixed at 200kt (p200), predict higher catches than in the other cases at moderate harvest rates, driven by higher mean recruitment level when spawning stock is above SSB_{break}. The more rapid decline in catches for the run where SSB_{break} is fixed at 200kt with increasing harvest rate is a result of an increasing proportion of replicates having SSB going below that value. In all cases the lower 5% percentile maximum of catches is at a lower harvest rate than that giving the maximum yield. The upper bounds of harvest target rate that results in less than 5% probability in the spawning stock biomass goes below the Btrigger (220kt) is in the range of ~0.21 to 0.23 in the all the scenarios tested (Figure 4.19).

The strong stabiliser in the HCR for Icelandic cod that goes abruptly off at $B_{trigger}$ reflects some desire for catch stability. Therefore, for the stakeholders in Iceland the rule is evaluated against $SSB < B_{trigger}$ and there is 5% probability to go below $B_{trigger}$=220 kt at HR_{target} = 0.23. This harvest rate does not take into account the stock assessment and implementation bias in recent decades.

Thus, the current harvest rule with a target rate of 0.20 achieves the objectives of having a low probability of going below $B_{trigger}$ of 220 kt **in the long term** and, achieves yields close to estimated maximum. There is relatively little gain in increasing the target rate to 0.22-0.23 in terms of long-term catches and maintaining the target at 0.20 also encompasses potential future bias of ~10-15%, in case of continued future assessment and/or implementation bias.

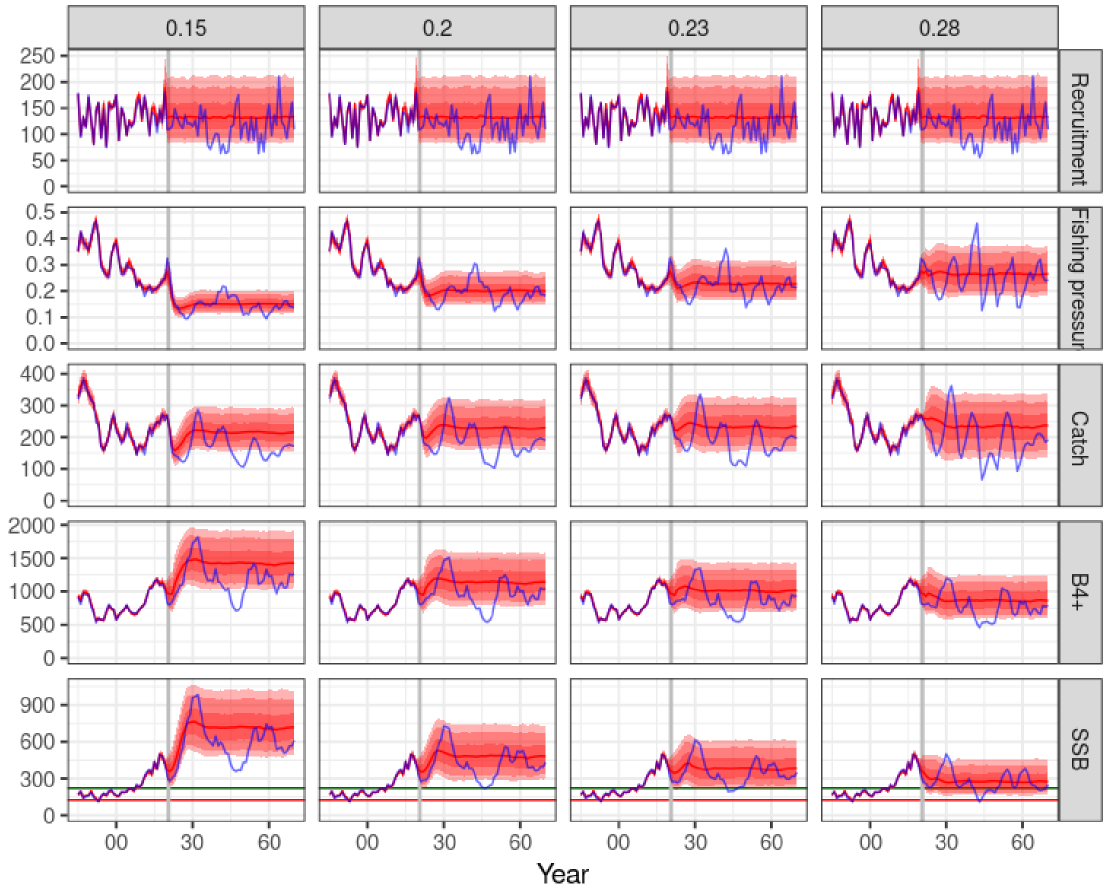

Figure 4.17: Assessment (from 1985 onwards) and projections of recruitment (millions), fishing pressure (realized harvest rate), catch (kt) and reference (B4+, kt) and spawning stock biomass (SSB, kt) by different harvest rate based on the base case. Red shades indicate 95%, 80% and 50% distribution range, the blue line one iteration. Horizontal lines refer to B_{lim} (red) and $B_{trigger}$ (green).

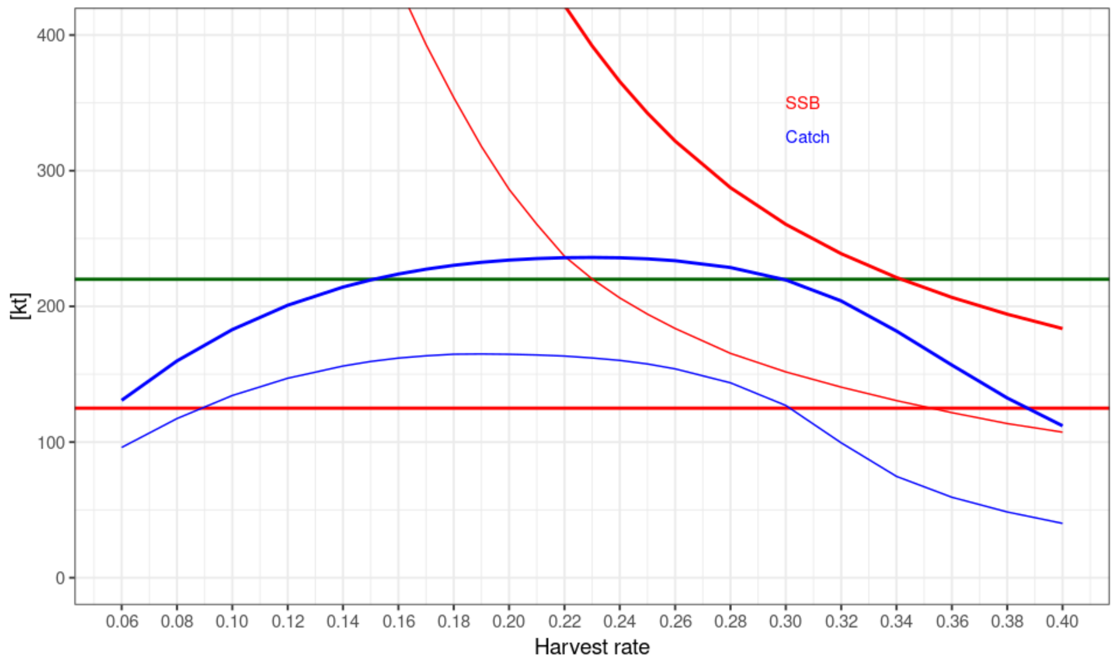

Figure 4.18: Base run: Mean catch (thick blue line) and 5% percentile catch (thin blue line), calculated on the basis of no Btrigger and mean spawning stock biomass (SSB - thick red line) and the 5% value (thin red line), calculated on the basis of Btrigger, as a function of harvest rate. Horizontal lines represent B_{lim} (red) and $B_{trigger}$ (green).

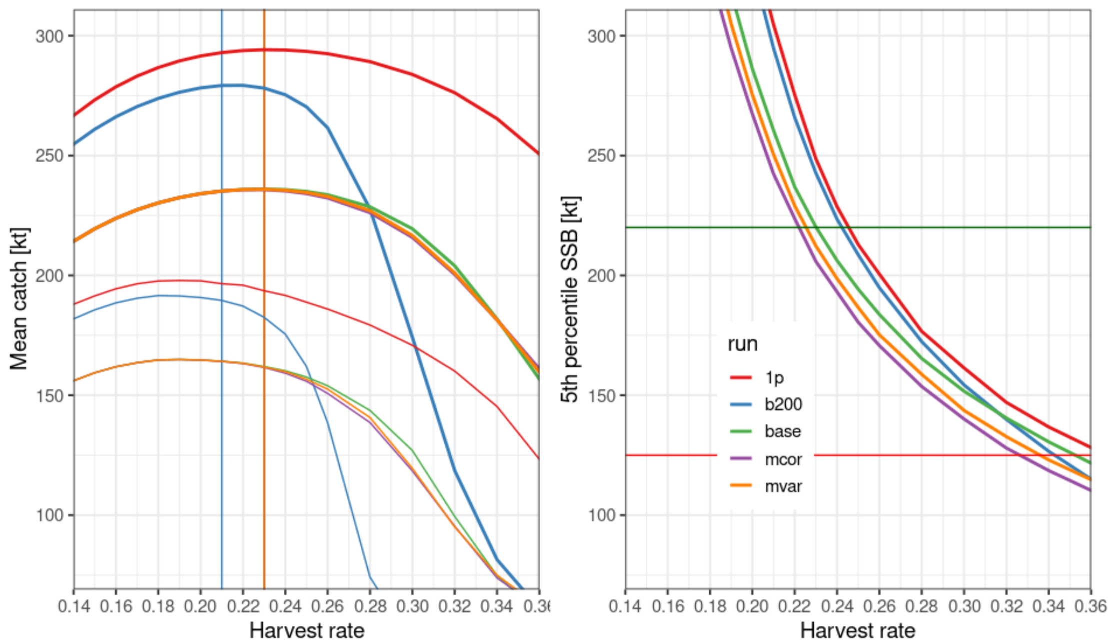

Figure 4.19: Mean catch (thick line) and 5% percentile catch (thin line), calculated on the basis of no Btrigger and the 5% percentile of the SSB, calculated on the basis of Btrigger, as a function of harvest rate for all scenarios tested. Horizontal lines represent B_{lim} (red) and $B_{trigger}$ (green). The scenarios are: '1p': no productivity change, 'p200': productivitivy change, SSB-break fixed at 200 kt, base: productivity change, SSB-break fixed at 120 kt, mcor: base case with correlated maturity, mvar: base case with stochastic maturity.

Summary of the median and the lower 5th percentile SSB and median catch for different management harvest rates (includes $B_{trigger}$ and catch stabilizer are presented in Table 4.1. 5th percentile, median and 95th percentile of catches, realized harvest rate, reference biomass and SSB are presented in Table 4.2.

Table 4.1: Cod in 5.a. Long term projected SSB and catches for harvest rates 0.20 to 0.23.

HR management target used in the HCR	Median SSB (in kt)	5th percentile SSB (in kt)	Median Catches (in kt)
0.20	492.409	286.240	230.512
0.21	455.392	260.529	231.904
0.22	421.887	236.961	232.975
0.23	391.972	220.225	233.649

Table 4.2: Cod in 5.a. Long term median, 5th and 95th percentiles of the projected reference biomass, SSB, realised harvest rate, and catches for harvest rates 0.20 to 0.23 and management Btrigger of 220 kt.

Target	var	5th percentile	Median	95th percentile
0.20	Catches (in kt)	163.031	230.540	317.844
0.21	Catches (in kt)	162.057	232.049	321.039
0.22	Catches (in kt)	159.984	233.211	323.981
0.23	Catches (in kt)	157.039	234.138	325.982
0.20	Realised harvest rate (HR)	0.150	0.201	0.273
0.21	Realised harvest rate (HR)	0.156	0.211	0.287
0.22	Realised harvest rate (HR)	0.161	0.220	0.301
0.23	Realised harvest rate (HR)	0.167	0.229	0.314
0.20	Reference biomass in kt	803.131	1142.545	1590.912
0.21	Reference biomass in kt	765.428	1095.480	1535.786
0.22	Reference biomass in kt	730.215	1052.635	1482.377
0.23	Reference biomass in kt	700.309	1012.535	1433.202
0.20	SSB (in kt)	289.617	479.395	739.862
0.21	SSB (in kt)	263.460	442.364	692.988
0.22	SSB (in kt)	240.260	408.759	649.737
0.23	SSB (in kt)	221.349	378.513	609.198

4.3.2 Short term risk

In ICES Technical Guidelines: Criteria for defining multi-annual plans as precautionary the following risk criterion are listed:

- **Risk1**: *Average* probability that SSB is below B_{lim}, where the average is taken across all iterations over some specified years.
- **Risk2**: *Proportion of iterations* where SSB is below B_{lim} *at least once* over some specified years.
- **Risk3**: *Maximum* by year probability that SSB is below B_{lim}, where the maximum is taken over some specified years. – This is ICES criterion.

Difference between Risk 1 and Risk 3 is usually in the first years of the simulations if the stock starts from a low level. Given the benchmark run, the stock in 2020 starts at 15% below the average level for HR=0.2 but the catches are relatively high and take, due to the stabiliser, some time to decrease. Therefore Risk 3 is higher than Risk 1 in the short-term and the maximum risk occurs after 3-4 years (Figure 4.20). The short-term risk of going below $B_{trigger}$ is here analogous to the results from the 2009 HCR simulation (see below) though in the latter case the short-term risk was higher because the median estimate of the SSB at the start of the simulation was much close to 220 kt.

The short-term risk is firstly associated with that the current TAC (fishing year 2020/2021), which is already set at ~257 kt, this being based on the 2020 assessment reference biomass estimate and the preceding TAC. Given the downward revision of the biomass the harvest rate is expected to be ~0.25. Secondly, what is used to start the assessment error sequence in 2020 for iteration j is:

$$log(\frac{\widehat{SSB_{2020}}}{SSB_{2020,j}})$$

$\widehat{SSB_{2020}}$ is a number given from a file (the average from the model). This method leads to the iterations with low value of stock starting with "overestimation" that will continue due to the autocorrelation of the assessment error. The opposite does apply to the iterations with high value of spawning stock, they start from underestimation.

The risk relative to $B_{lim} = B_{loss}$ is much lower than 0.05 irrespective of risk1 or risk3 criterions and the time periods considered.

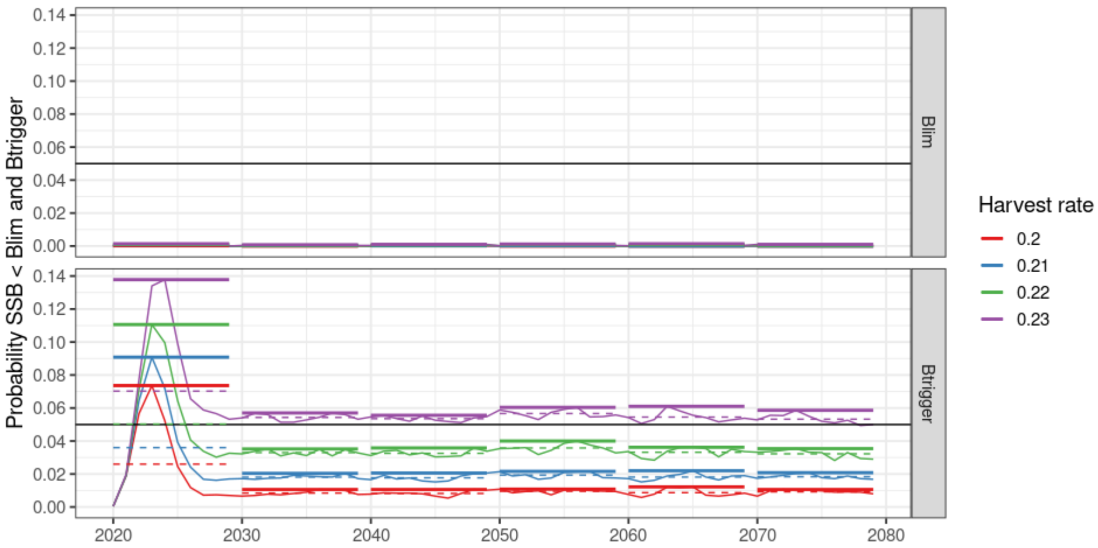

Figure 4.20: Risk 1 (horizontal full line over each future decade) and 3 (horizontal dashed line over each future decade) based on the benchmark assessment. Upper panel: Risk to B_{lim}, lower panel: Risk to $B_{trigger}$.

4.4 Robustness of the biomass rule to alternative selection patterns

Muppet estimates selection by age in 4 periods 1955-1975, 1976-1993, 1994-2006 and 2007-onwards. The selection is only constrained by being the same for oldest three age groups (ages 12 to 14). The surveys are on the other hand constrained to have the same q for ages 10 and older. The selection patterns from Muppet are relatively similar between periods, the selection in the last period is caused by the surveys affecting ages 10-13 (before 2010 there were relatively few age 10 and older in the surveys, mostly noise).

Selection patterns from SAM are usually more dome-shaped than estimated in muppet, but in most recent years the selection obtained from SAM are much less dome shaped, mostly controlled by the surveys (see Figure 3.37). To understand the effect of selection pattern on yield it is first necessary to look at the predicted virgin biomass in each age group assuming $Z = M = 0.2$ (Figure 4.21). Apparently, it is best to avoid catching fish younger than age 8 and we should also try to catch all 14 years old fish (no plus group in this model). A dome-shaped selection pattern would lead to a loss of catch but does still give us virtual SSB.

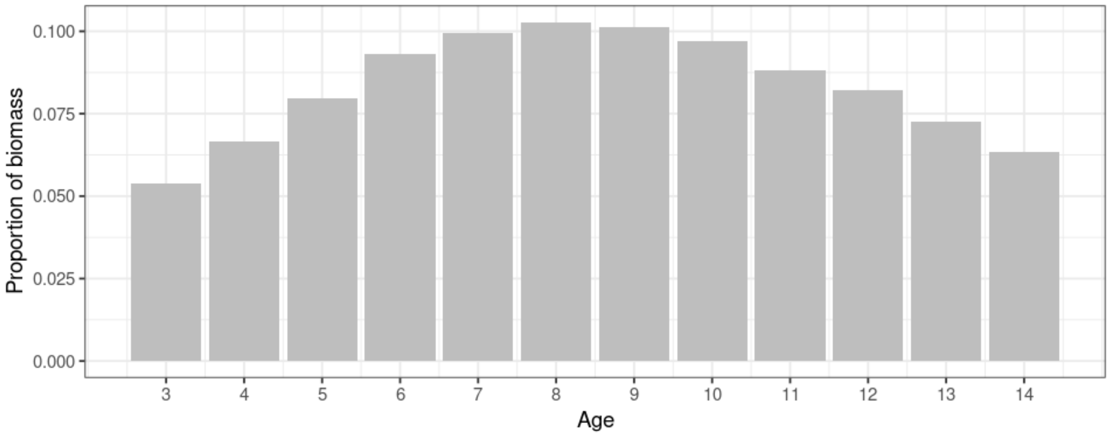

Figure 4.21: Predicted composition of virgin biomass assuming natural mortality of 0.2.

Estimated average selection patterns from SAM and muppet are shown in Figure 4.22. All the models lead to the same 5th percentile of spawning stock. There is some difference in yield between different periods in muppet but all the SAM runs get lower catches even the run $sam2019$ where the selection pattern is not very dome-shaped.

Investigation of the results indicated that a large part of the difference between SAM and muppet was caused by variability in selection, but selectivity at age involves estimated parameters in muppet and each future iteration has its own selectivity. Investigation shows that fixing the future selectivity at the average from muppet moves the results in terms of catch closer to the SAM values, i.e lower median yield.

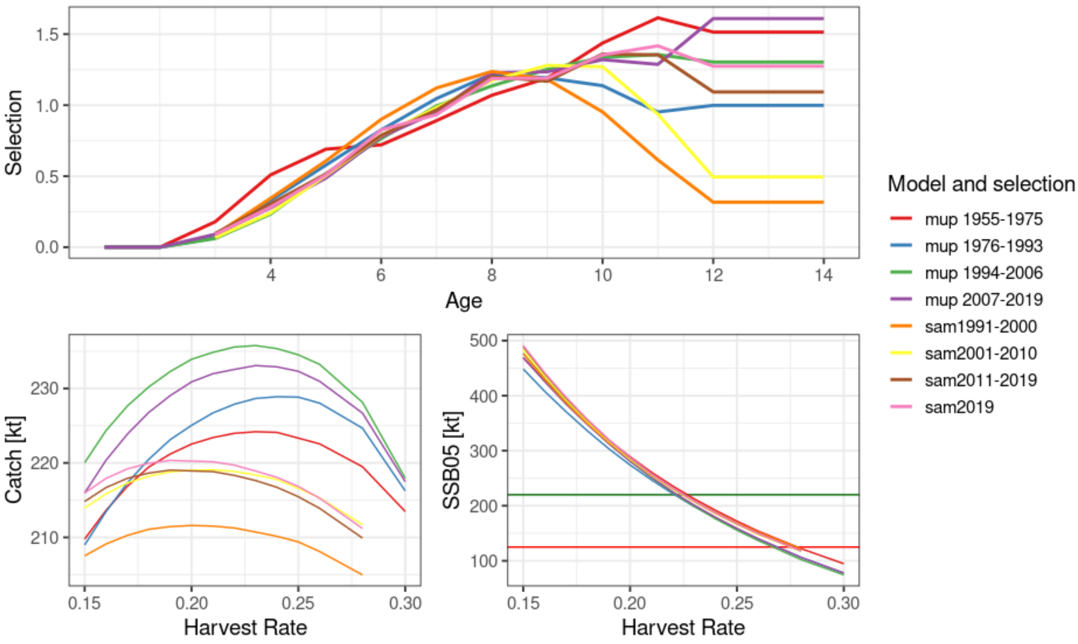

Figure 4.22: Robustness of the 5th percentile of SSB to selection different selection patterns estimated from muppet ('mup') and SAM. Horizontal lines in the SSB plot refer to Blim (red) and MGT Btrigger (green).

There is some gain with fishing cod at older ages, the question is really what selection pattern is practical, for example due to mixed fisheries considerations and seasonality of fisheries. Operating trawls with large mesh size leads to the small fish escaping through meshes, possibly with some mortality. Also, it is not clear at what age M starts to increase but for individual fish that is both temperature and size dependent. For illustration, results of 5 idealised knife-edge selection patterns are shown (Figure 4.23).

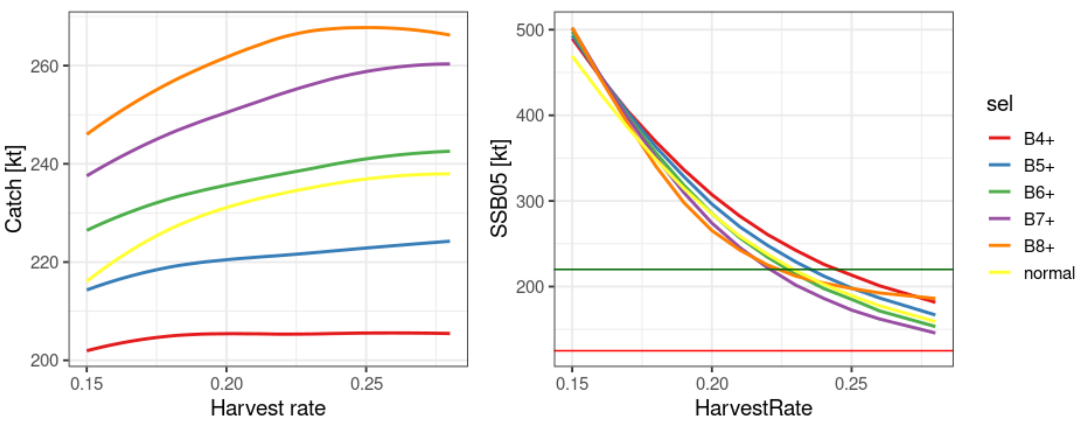

Figure 4.23: Median yield and fifth percentile of SSB for knife edge selection patterns 4+, 5+ …8+. Results using estimated selection shown for comparison. Horizontal lines in the SSB plot refer to Blim (red) and MGT Btrigger (green).

In principle there is a problem with knife-edge rule like B4+ if there is, in fact, a dome shaped selectivity pattern. This problem is theoretically less if an F-based rule on the realised selection pattern is used. In practice, accepting a dome-shaped selection pattern is not trivial so the F rules do not in practice, tackle this kind of problem.

Estimated dome shaped selection pattern is caused by the Z signal in catch in number at age but Z is higher than expected from flat selection pattern. Three factors can explain this high Z, dome-shaped selection, increased M of older ages and negative bias in age readings. If a stock assessment is based on catch in number plus a few surveys, at least one of fleets must be assumed to have flat selection at oldest age groups, the problem can also be solved by a plus group.

4.5 Comparison with the simulation done in 2009

In the simulations done in 2009 (ICES (2010)), that were the basis for the reduction in target harvest rate from 25% to 20%, it was expected that the reference stock and particularly the spawning stock would increase (Figure 4.24), the latter in part because of expected increase in survival of older (age 7+) cod. The harvest rule was tested, among other things, assuming the mean future recruitment would reflect that observed since 1985 (box-distribution bounded by the minimum and maximum estimated recruitment) as well as using the recent low weights observed in the stock in the most recent years (then 2005-2009 - see Figure 4.2). These simulations were otherwise built on the same framework (model and code) as is currently used.

Comparing the simulation done in 2009 with the benchmark assessment (Figure 4.24) show that the "realized" recruitment is in the upper range of that assumed, and consequently also the reference biomass and the catches. Fishing pressure has been within the range expected, but according to the current assessment is increasing in the last three years. It should be noted that the expected spawning stock biomass response in 2009 was at that time an extrapolation of any the recent historical values. Although this comparison is by no means a validity test it gives some credence to the framework used.

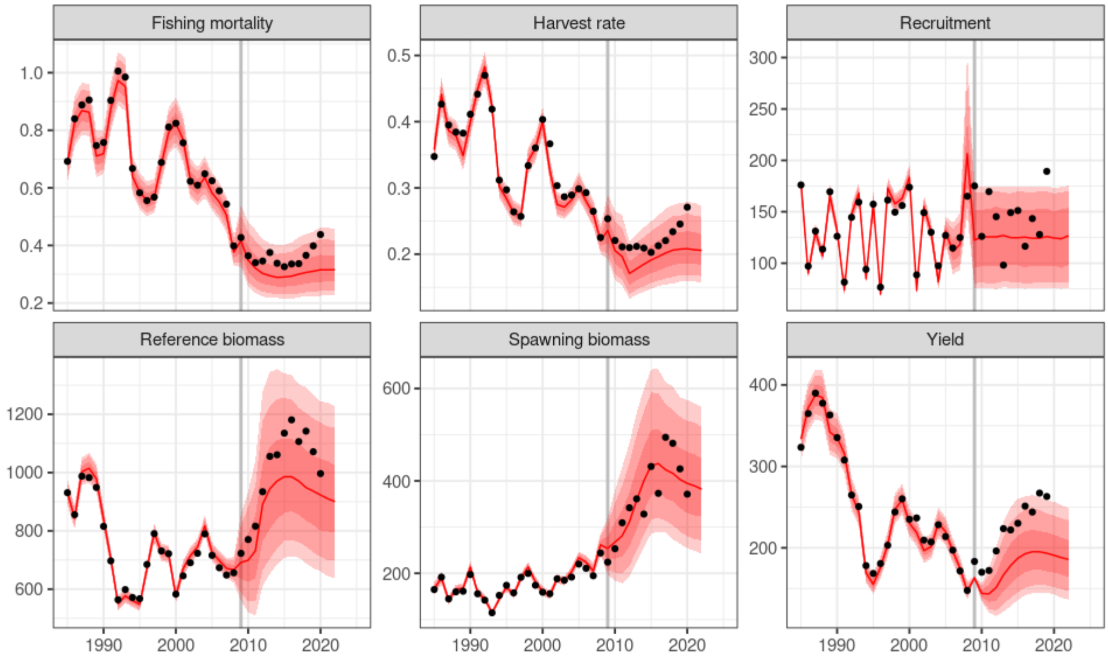

Figure 4.24: Comparison of distribution of key metrics of the simulation done in 2009 if a 20% harvest rule was followed (red ribbons) with the 2020 assessment (black dots).

4.6 ICES reference points

4.6.1 PA points

B_{lim}

- The current $B_{lim} = 125$ kt based on $B_{loss} = 125$ kt
- Based on the new assessment the $B_{loss} = 115$ kt. VPA gives $B_{loss} = 120kt$ and SAM $B_{loss} = 125kt$.
- Suggested to keep the 125 kt, and the basis and source then B_{loss} from ICES 2010a.

HR_{lim}

- Derived by using no stabiliser, no assessment error and no Btrigger.
- Currently $F_{lim} = 0.74$.
- New calculation based on harvest rate is: $HR_{lim} = 0.35$.

B_{pa}

- Currently $B_{pa} = 160$ kt, the basis being $B_{pa} = B_{lim}e^{(1.645\sigma_B)}$ where $\sigma_B = 0.15$.
- No changes proposed

HR_{pa}

- Currently $F_{pa} = 0.58$ is defined. This was derived from F_{lim}, a procedure that is no longer the basis in the guidelines.
- The current guideline state "The fishing mortality including the advice rule that, **if applied as a target in the ICES MSY advice rule (AR)** would lead to SSB ≥ Blim with a 95% probability (also known as Fp05)." This is understood to be such that the ICES $MSYB_{trigger} = 265$ kt (see below) is applied and **assessment error is included**. This results in a $HR_{pa} = 0.39$. $HR_{pa} > HR_{lim}$ because latter is based on no trigger and error, while the former is.
- At harvest rates exceeding 0.25, more and more of the replicas will be below ICES $MSYB_{trigger}$ so $HR_{pa} = 0.39$ imply an average realized harvest rate of ~0.30.

4.6.2 ICES MSY points

Currently the ICES MSY reference points are set using the same the management HCR rule. It is likely that these values were at the time just quickly adopted by ICES. In addition, since set, the guidelines may have changed. The least confusing approach, considering the recipients (read: stakeholders) is to do the same, again. What follows are the derivations using the most recent guidelines.

ICES HR_{msy}

- Currently ICES $HR_{msy} = 0.20$.
- New calculation is $HR_{msy} = 0.24$
  - 2.5% assessment bias (analytical retrospective): 0.23
  - 7.0% assessment bias (historical retrospective): 0.22

ICES MSY $B_{trigger}$

In the guidelines it is stated: *"In the ICES MSY approach, MSY Btrigger is defined as the* 5^{th} *percentile on the distribution of SSB when fishing at FMSY. This calculation* ***does not include assessment/advice error****, but includes annual stochasticity in population parameters and fishery selectivity. When a stock declines below MSY Btrigger, this triggers advice for a reduced fishing mortality compared to FMSY."*

The optimum harvest rate in the absence of assessment error is $HR_{msy} = 0.24$ (0.23 including assessment error) and this results in ICES $MSYB_{trigger} = 265kt$ (same value including assessment error).

4.6.3 Additional considerations

The principle behind the ICES $B_{trigger}$ is somewhat sensible, one does not want to be too often below the $B_{trigger}$. The other side of the coin is that the stabilizer is not included in compiling ICES HR_{msy} and $B_{trigger}$. The stabilizer has little effect on HR_{msy} and average spawning stock but SSB_{05} reduces to 210 kt if stabilizer is included. This is expected behaviour since a stabilizer should lead to less variability in catches but should increases variability in SSB. Another problem with the ICES approach is that if ICES HR_{msy} would have been estimated higher the $B_{trigger}$ would have been lower, contradicting the normal behaviour of HCR where $B_{trigger}$ and HR_{target} are negatively correlated.

What drives the HR_{msy} obtained here is really the yield-per-recruit, if we run with stabilizer and without trigger SSB_{05}, is at B_{loss} (estimated SSB_{break}) at $HR = 0.265$. At higher harvest rates the stock-recruitment function starts kicking in and yield drops fast. Here the stabilizer starts to have effect earlier as the lower quantiles of SSB are lower. Without the stabiliser and without trigger SSB_{05} is at B_{loss} at $HR = 0.3$. If we assume a Ricker function the peak of that function will dominate where HR_{msy} is but the Ricker function does not fit the data as well as the Hockey stick (not a proof of anything, both are wrong)

The problem is that the two factors, yield-per-recruit and SSB-recruitment relationship, are giving maximums that work a little against each other, $MSYB_{trigger}$ is obtained from the yield-per-recruit maximum but used for precautionary point (HR_{pa} or F_{05}) that is more linked to the SSB-recruitment relationship. The result in a high HR_{pa} value but realized harvest rate being lower (~0.30).

Additional factor in the derivation of the ICES reference points that requires clarifications is why advice error is sometimes included and sometimes not.

5 References

Björnsson, Eyþór. 2004. "Olíunotkun íslenska fiskiskipaflotans og losun gróðurhúsalofttegunda frá honum." *Háskólinn á Akureyri, Auðlindadeild*.

Guðmundsson, Guðmundur. 1992. "Fiskveiðistjórnun." *Fjármálatíðindi*.

ICES. 2010. "Report of the Ad hoc Group on Icelandic Cod HCR Evaluation (AGICOD." *ICES CM 2009/ACOM:56. 89 Pp.* Vol. ICES CM 2009/ACOM:56. 89 pp. ICES. https://doi.org/10.17895/ices.pub.5279.

Jonsson, S. T., and E. Hjorleifsson. 2000. "Stock Assessment Bias and Variation Analyzed Retrospectively and Introducing the PA-Residual." *ICES C.M. 2000 / X:9*.

NN. 1994. "Hagkvæm nýting fiskistofna. Vinnuhópur um nýtingu fiskistofna." *Skýrsla Til Sjávarútvegs-ráðherra*. https://fishvice.hafro.is/lib/exe/fetch.php/icod:hcr:icodhcr1994report.pdf.

———. 2004. "Aflaregla fyrir þorskveiðar á Íslandsmiðum. Skýrsla nefndar um langtímanýtingu fiskistofna. Vinnuhópur um nýtingu fiskistofna." *Skýrsla Til Sjávarútvegsráðherra*. https://fishvice.hafro.is/lib/exe/fetch.php/icod:hcr:nefndaralit2004lokaeint.doc.

SCHOPKA, S. 1994. "Fluctuations in the Cod Stock Off Iceland During the Twentieth Century in Relation to Changes in the Fisheries and Environment." *ICES Mar. Sci. Symp.* 198: 175–93. https://ci.nii.ac.jp/naid/10022009197/en/.

6 Reviewers' comments

Reviewers' report

Malcolm Haddon (University of Tasmania), Margarita Rincón Hidalgo (Spanish Oceanographic Institute IEO) and Bjarte Bogstad (Institute of Marine Research, Norway).

Overall conclusion

All the work described in the report as well as the presentations and clarifications during the workshop led to the conclusion that the settings of the stock assessment model, the data used and the model fit are adequate, and that results generated by the assessment model can be used for providing fisheries advice as an ICES Category 1 stock (stocks with quantitative assessments). Different issues identified for future consideration are noted below in the "Stock assessment comments" section.

The estimation of the precautionary and MSY reference points has been conducted in accordance with the ICES guidelines.

Regarding the MSE framework, the "short-cut" approach to simulating future assessment modelling was carefully done accounting appropriately for different sources of uncertainty, mainly on the stock-recruitment relationship specification and variability based on the historical assessment performance, showing also individual simulated trajectories. Only management plans with the same type of formulation as the existing one were evaluated, the management plans proposed were tested against different assumptions including a change in productivity and different selection patterns, with no remarkable differences, showing the robustness of the modelling framework.

The existing harvest control rule has a harvest rate (HR) applied to the biomass of age 4 and older fish (exploitable biomass, B4+) set to HRMGT=0.20 and a trigger point for SSB at MGT Btrigger= 220 kt, with the harvest rate reduced proportionally when SSB is estimated to be below the trigger. A stabilizer is applied when SSB> =MGT Btrigger= 220 kt, so that the advised catch then is calculated as the average of HR* B4+ and the previous year's TAC.

This rule is precautionary (it leads to less than 5% probability of SSB < Btrigger in all years) and, when the harvest rate is below HRMSY=0.22, the long-term simulations indicate the same precautionary level; this is because the resulting curve of average catch in equilibrium versus harvest rate (Figure 4.18) is flat-topped, with the average long-term catch at HRMGT=0.20 being slightly below that corresponding to HRMSY=0.22. Therefore, the proposed harvest control rule can be considered to be in conformity with the MSY approach.

Comments on stock assessment

Overall the performance of the muppet stock assessment model chosen was good. Although none of the model configurations tested were able to fit "particularly well" some rapid changes in the survey data (Figure 3.24), the diagnostics did not show any major issues in the model's ability to follow the average patterns of the main processes determining the productivity of the stock with a reasonable retrospective performance (albeit a slight bias in reference biomass estimates ~2.4%, Figure 3.31). However, a few points were still suggested by the reviewers for future exploration:

- Given the changes in the proportion of total catch taken by different fishing gears since the 1980s, it is suggested that additional time-blocks of alternative selectivity patterns should be examined to determine, at least, to what extent such changes may contribute to uncertainty in model outcomes.

- Explore options for including multiple fleets to represent the alternative fishing gears. This would allow for varying the fully selected fishing mortality rates between gears.

Comments on reference points

The biomass reference points Blim and Bpa remain as they were in the latest HCR cod evaluation (ICES 2010a) following previous ICES guidelines for reference points calculations.

Taking into account the current differences in legislation and management between the countries receiving advice from ICES, it will be advantageous to simplify or reduce the number of reference points to no longer require all RP to be calculated in all circumstances.

An additional factor in the derivation of the ICES reference points that requires clarifications is why advice error is sometimes included and sometimes not.

General comments

Regarding the process for this review, the following comments apply:

- Most of the reviewers' comments during the MSE workshop were clarifications or referred to aspects that could be explored in future work rather than being urgent matters that needed to be resolved or changed immediately.
- The introductory meeting (the week previous to the workshop) and main documents presented were useful as they were fairly complete.
- Having the WDs and the reproducible code that generate them available in a transparent. framework was useful. This should also help with easing tasks such as including this into the TAF ICES framework.
- Having direct access to the software used in all the analyses was very helpful in allowing the reviewers clarity over the dynamics used in the modelling.
- Additional time at the beginning to cover the process for running the model and generating results, and some sort of introduction to the fishery might have helped with both uncovering potential issues and helping with understanding the model.
- The time allotted for the work was relatively short, especially given the expanded tasks of evaluating the MSE.
- It is recommended that requests from managers on which HCRs to be tested should be made more explicit, rather than implicit as they were in this case.

Annex 1: List of participants

Name	Institute	Country (of institute)	Email
Höskuldur Björnsson	Marine and Freshwater Research Institute	Iceland	Hoskuldur.bjornsson@hafogvatn.is
Bjarte Bogstad (Chair)	Institute of Marine Research	Norway	Bjarte.bogstad@hi.no
Malcolm Haddon	University of Tasmania	Australia	malcolm@haddon.net.au
Margarita Rincón Hidalgo	Instituto Español de Oceanografía	Spain	margarita.rincon@ieo.es
Einar Hjörleifsson	Marine and Freshwater Research Institute	Iceland	Einar.hjorleifsson@hafogvatn.is
David Miller	ICES	Denmark	David.miller@ices.dk

Additional participants in parts of the online meetings

Mollie Brooks	DTU Aqua, National Institute of Aquatic Resources	Denmark	molbr@aqua.dtu.dk
Daniel Howell	Institute of Marine Research	Norway	daniel.howell@hi.no
Bjarki Elvarsson	Marine and Freshwater Research Institute	Iceland	bjarki.elvarsson@hafogvatn.is
Polina Levontin	Imperial College London	UK	polina.levontin02@imperial.ac.uk
Warsha Singh	Marine and Freshwater Research Institute	Iceland	warsha.singh@hafogvatn.is

Annex 2: Special request to ICES from Iceland on re-evaluation of the management plan for Icelandic cod

ICES - International Council
for the Exploration of the Sea

Att: Lotte Worsøe Clausen Mark Dickey-Collas
H.C. Andersens Boulevard 44-46
DK-1553 Copenhagen V
DENMARK

ATVINNUVEGA-OG
NÝSKÖPUNARRÁÐUNEYTIÐ

Ministry of Industries and Innovation

Skúlagötu 4 101 Reykjavík Iceland
tel.: + (354) 545 9700 postur@anr.is
anr.is

Reykjavík November 9, 2020
Reference: ANR19070057/02.03.07

Subject: Subject: Re-evaluation of the management plan for Icelandic cod.

Reference is made to the Memorandum of Understanding between Iceland and ICES, signed 1.12. 2019. The Government of Iceland is in the process of re-evaluating the management plan for the Icelandic cod stock (cod.27.5a). The management strategy for Icelandic cod is to maintain the exploitation rate at the rate which is consistent with the precautionary approach and that generates maximum sustainable yield (MSY) in the long term.

Part of the management plan is the adoption of harvest control rule (HCR) for setting annual total allowable catch (TAC). The HCR adopted should be precautionary and in accordance with the ICES MSY approach. The current management plan for cod was first evaluated by ICES before the 2009/2010 fishing year and was re-evaluated in 2015 and found to be consistent with the precautionary approach and in conformity with the ICES MSY-framework.

In a letter from the Ministry dated on the 20th of November 2019, ICES was informed that the Minister of Fisheries and Agriculture had appointed a working group to review the management plan and HCR for cod in Icelandic waters. The work was expected to finish by the end of May 2020. Due to disruptions by the COVID-19 pandemic the work was delayed but is expected to finish in December. The main outcome of this work will be a proposal on a HCR for cod, either in the form of the current HCR or some variants of the rule. Technical documentation of the proposed HCR by the aforementioned working group will be produced by national experts at the Marine and Freshwater Research Institute and made available to ICES before the 15th of February 2021.

The Government of Iceland requests ICES to evaluate whether the proposed harvest control rule or rules are in accordance with its objectives, given current ICES definition of reference points or any re-evaluation of those points that may occur in the process. Additionally, the evaluation should also include review of input data and the applied assessment methodology for cod (Benchmark). It is expected that the ICES advice for the 2021/2022 fishing year for Icelandic cod (cod.27.5a) be based on the above-mentioned HCR.

On behalf of the Minister of Fisheries and Agriculture

Áslaug Eir Hólmgeirsd.

Áslaug Eir Hólmgeirsdóttir

Þorsteinn Sigurðsson

Cc: Marine and Freshwater Research Institute:

Annex 3: Summary Template for HCR modelling

Stock: Cod 5.a

Background		
Motive/ initiaitve/ background.	5 year reevaluation of the HCR, check if current harvest rate is valid.	
Main objectives	In conformity with the precautionary approach and maximum sustainable yield.	
Formal framework	ICES on request from Iceland.	
Who did the evaluation work	WKICECOD 2021	
Method		
Software Name, brief outline include ref. or documentation	Software: Assessment and simulation done within AD model builder. Downstream analysis and documentation done in R. Name, brief outline: Muppet. Statistical catch-at-age separable model, both the historical assessment and future simulations are done in the same framework. Assessment based on catch age 3-14 and two survey indices age 1-14/3-13. Shortcut approach (no assessment in the loop), historical retrospective errors applied to the true reference biomass (B_{4+}) and SSB when dictating future TAC. Reference: https://github.com/Hafro/Muppet_HCR	
Type of stock	Long lived, demersal, valuable	
Knowledge base *	Analytical age based assessment.	
Type of regulation	TAC / ITQ.	
Operating model conditioning		
	Function, source of data	Stochastic? - how (distribution, source of variability)
Recruitment	Hockey-stick fitted to SR pairs 1955–2019. Alternatives explored.	Lognormal distribution, SSB_{break}, CV and one productivity change (downward) estimated internally, autocorrelation applied externally based on estimates from model. Alternative hockey sticks explored.
Growth & maturity	Average from 2011-2020	Weights: Deviations lognormal applied to all age groups within a year. CV and autocorrelation of the deviations based on average of estimated values for ages 4-10. Maturity: No stochasticity. Alternatives explored, i.e stochasticity and correlated with mean weight at age
Natural mortality	Constant for all ages and year.	No
Selectivity	From the last separable period of the assessment (starting 2007)	Eestimates from each historical iteration carried forward.

Initial stock numbers	From the historical assessment, 2020 starting values from each iteration.	According to variance - covariance matrix from assessment (inverse Hessian)
Decision basis **	Base case: TAC for next fishing year (starting September 1st in the assessment year) is a function of TAC in current fishing year (weight 0.5) and estimated B_{4+} in the beginning of the assessment year (weight 0.5). If SSB in the assessment year is estimated below management $B_{trigger}$, harvest rate is reduced and the stabiliser removed. The main characteristics is that TAC is only a function of B_{4+} in the assessment year so in principle no short term prediction is required.	
Number of iterations	2000	
Projection time	60 years	
Observation and implementation models		
Type of noise	Lognormal noise applied to the true reference biomass (B_{4+}) and SSB in the assessment year based on CV (0.12) and autocorrelation (0.52) estimated from the historical (empirical) assessment performance 1970-2015. Analytical retrospective patterns indicate lower CV than used here or ~0.05.	
***** Comparison with ordinary assessment?**	Future retrospective pattern (CV and autocorrelation) the same as the historical patterns.	
Projection: If yes - how?	No projection needed to decide the TAC.	
Projection: Deviations from WG practice?	Not applicable.	
Implementation	No implementation error assumed, effect evaluated qualitatively externally.	
Harvest rule		
Harvest rule design	If $SSB_y < B_{trigger}$: $HR_{y/y+1} = HR_{std} \times SSB_y / SSB_{trigger}$ If $SSB_y >= B_{trigger}$: $HR_{y/y+1} = HR_{std}$	
Stabilizers	If $SSB_y >= B_{trigger}$, final TAC is half the value of above and half of the TAC in the present fishing year. Not applied when $SSB_y < SSB_{trigger}$.	
Duration of decisions	Annual	
Revision clause	Normally evaluated every 5 years.	
Presentation of results		
Interest parameters	Risk, average and stability of catch, ssb, reference biomass, recruitment and harvest rate (the mean and the 10th, 50th, and 90th percentiles).	
****** Risk type and time interval**	Type 1 and 3, for decades 2020-2029 to 2070-2079. Short term considerations are most important here.	
Precautionary risk level	5%	
Experiences and comments		
Review, acceptance:	Accepted by review group, implemented from 2012 onward.	
Experiences and comments	The evaluation is done relative to $B_{trigger}$, the B_{lim} is based on B_{loss}. The reason is related to productivity change (≈ 35% decrease in recruitment) that seems to have occurred around 1985. SSB_{break} is estimated at B_{loss} (~120kt) that is used as B_{lim} both with productivity change included and not included. In earlier HCR evaluations if productivity change was not estimated, SSB_{break} was 2009 and before estimated at around 220 kt (that is current $B_{trigger}$), 171 kt in 2015 and at B_{loss} in the 2021 HCR evaluations . With productivity estimated SSB_{break} has always been estimated at B_{loss}. Comparing SSB_{break} at B_{loss} vs $B_{trigger}$ the values are -1998.28 and -1991.71	

	with productivity change but -1986.916 and -1985.24 with no productivity change. Evidential support for the productivity change is therefore strong, the objective function changes by -12 for one estimated parameter that is a large change for $SSB - Rec$ parameters. Not including productivity change does not support $SSB_{break} = B_{loss}$ as strongly as when productivity change is estimated. Autocorrelation is estimated very low (~0.05) when productivity change is estimated but 0.2/0.3 when it is not estimated. Productivity change and $SSB_{break} = B_{loss}$ does therefore seem like the most plausible settings but not productivity change and $SSB_{break} = 220$ kt was closer in terms of objective function in evaluations 2009 and earlier (2015 was in between).

FOOTNOTES:
* **Knowledge base**: This is the information that will be available about the state of the stock, in particular whether there is an assessment or not. If it is something else, please specify.

** **Decision basis**: This is the measure that determines the exploitation in the harvest rule. For example, SSB at the start of the TAC year, TSB in the last assessment year,.

*** **Comparison with ordinary assessment?** This is to indicate whether there has been attempts to verify that the performance of the assessment in the model is similar to that experienced by the WG, for example with respect to retrospective problems and inconsistencies.

**** **Risk types:**

- **Risk1** = *average* probability that SSB is below Blim, where the average is taken across the ny years.
- **Risk2** = probability that SSB is below Blim *at least once* during the ny years.
- **Risk3** = *maximum* probability that SSB is below Blim, where the maximum is taken over the ny years.

If your definition of risk does not fit any of these, please explain.

Annex 4: Investigation of harvest rates taking cost of fishing and price into account.

10 Investigation of harvest rates taking cost of fishing and price into account[1].

10.1 Introduction

In the evaluations of Harvest control rule for Icelandic cod 1993-1994 (NN (1994)) and 2003-2004 (NN (2004)) the goal was to maximize current value of the revenue of the fisheries instead of the catch, the current value was based on 5% required return. It can be argued that 5% required return is too much for a renewable resource that should never be depleted. In 1994 a major reduction in effort and catch was unavoidable and how fast effort should be reduced was one of the problems investigated. The result was that looking only at maximum revenue closing the fisheries for few years was the best option. Number of important factors, like loss of trained employees and lost markets were then ignored. The economic models are described in the reports from 1994 and 2004, they are similar but not identical. Among factors included are.

1. Effect of supply on price of cod. The work was based on 7% lower price with doubling of cod catch.

2. Cost per unit effort is fixed but catch per unit effort increases with size of fishable stock in the power 0.7 so doubling of fishable stock leads to 62% increase in CPUE. Fishable biomass is defined in the traditional way i.e from average number of fish in the year, estimated selection pattern of the fleet and mean weight at age in the catches. $Fb_y = \sum_a N_{y,a} \frac{F_{ya}}{Z_{ya}} Sel_a cW_{y,a}(1-e^{-Z_{ya}})$ where Sel_a is average selection pattern of the fleet (Figure 10.11).

3. On the average price of larger cod is higher, there is though some interannual variability in the trend. Equation based on data from fishmarkets in 1997 that was used in the work 2003-2004 is $p = 56.93 + 7.91W$ (NN (2004)) or 15% increase in price for each kg in average weight.

4. In 1994 reduction of capelin and shrimp (*Pandalus borealis*) catch with increased cod stock was taken into account, not in subsequent analysis. The result from 1994 was that inclusion of predation was not necessary and the predator-prey relationships between cod and capelin and shrimps are complicated and unpredictable. When the evaluations in 1994 were done catch of shrimp in term of values was ~30% of value of cod catch but the shrimp stocks in Iceland collapsed in the late 1990's due to cod predation, more a spatial overlap issue than large cod stock.

What has most weight in the economic model is reduced cost of fisheries with large stock. What is calculated in the economic model is social revenue so wages are not included as cost. The 1994

[1] Annex 4 was named Section 10 in the draft report, hence the numbering of subsections, tables and figures in this Annex.

report did though discuss opportunity cost i.e the value that fishermen could create if they were not fishing. Before 1994 other authors had demonstrated the economic benefit of reducing fishing effort (Guðmundsson (1992)).

What is then left is the cost of vessels, fuel and gear that is mostly import cost. The cost of a operating a vessel is to large extent independent of use once the vessel has been acquired. Therefore, major reduction in cost will not be reached until the number of vessels is reduced in accordance with reduced effort. Cost of oil and bait decreases immediately after reduction in effort and cost of gear relatively quickly.

Mixed fisheries can lead to reduced economy of the fisheries compared to what it can mostly be. Cod is a rather easily caught species and large part of many fishing trips are spent on other species even though cod accounts for most of the value. This can be an indicator that harvest ratio should be reduced for species that are more difficult to catch. That kind of action would though not necessarily solve the problem. Even though the ratio between species were close to correct taking into account economy of the fisheries that does not have to apply to individual companies. This has been a problem for many years regarding haddock. The problem is related to change in spatial distribution of haddock where increased proportion of the stock is in fishing areas north of Iceland but proportion of the quota of the fleet operating in the north does not change as rapidly. Around the turn of the century effort by longliner increased but species composition in the catch of longliners is different from what it is in bottom trawl and gillnets that are the gears that had a decreasing share. Many companies have not yet changed their ITQ portfolio to match longline fisheries.

In recent years increased proportion of the cod catch has been exported fresh. This fish is preferably caught in the last 2 days before landing. This creates certain by-catch problem in the earlier part of each fishing trip but does also show the importance of moderate harvest rate that allows to catch the fish in shorter time but the price of recently caught fish probably higher than for fish caught 4-7 days ago. Most of the considerations here apply to bottom trawl fisheries but in recent years 43% of the cod catch has been caught by bottom trawl but 54% of the total demersal catch. Here it is assumed that the revenue from fisheries using other gear is similar to bottom trawl fisheries.

Effects on the environment will be of major concern in coming years. Among factors that will be investigated are:

- Effect of fishing gear on bottom substrate.
- Effect of fishing gear on fish that is not caught, discards, mesh penetration, fish slipping from the hooks etc.
- CO_2 footprint (fuel consumption).
- Garbage left, longlines, ghostnets etc.

Effect of all those factors decrease with decreasing harvest rate, even though catch does not decrease. The effects are most apparent in bottom trawl fishery. In some cases, the area trawled does also shrink considerably. Bottom trawl fisheries have today negative stamp due to negative effects on the environment. High CPUE reduces those negative effects and are therefore important for acceptance of trawl fisheries. Although other gear can partly replace bottom trawl it will still be required in a fleet that utilises demersal fish stocks in Icelandic waters.

10.2 Mixed fisheries.

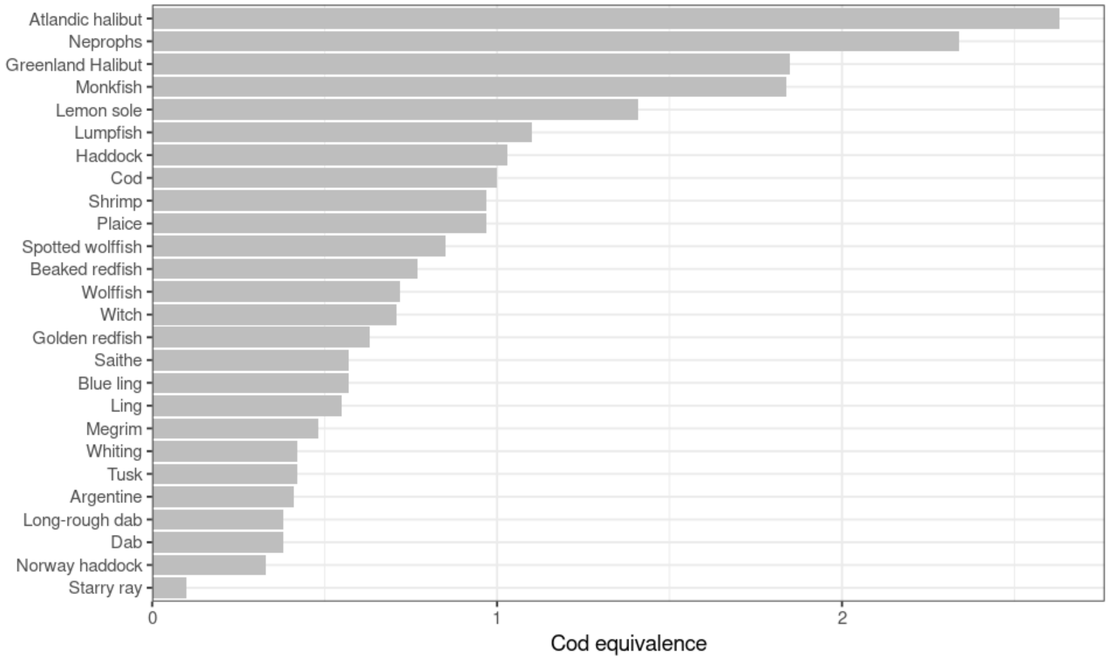

Figure 10.1: Cod equivalence coefficients for some species

The demersal fleet catches more than 30 species. TAC is allocated for the most important species. It can be transferred between species to rather limited extent and the relative value of species in transfer is based on cod equivalence coefficient of the species (Figure 10.1). This value is based on the price per kg of the species but does not take into account the cost of catching the species that varies between species. Saithe is an example of species that is relatively difficult to catch and the TAC has not been caught in recent years. The reason is that saithe is more pelagic than other demersal fish stocks and therefore less available for demersal gear. It does not help that the price is not high. Harvest rate of saithe has never been high even before fisheries were limited by quotas. Earlier, third part of the saithe catch was taken by gillnets but gillnet fisheries for saithe stopped for more than 25 years ago, in the period where number of longliners increased. Longliners catch saithe to very limited extent so increased longline fisheries lead to reduced saithe fisheries.

Another example is deepwater shrimp (*Pandalus borealis*) where fuel cost as proportion of the value of catch is four times higher than in demersal fisheries (Björnsson (2004)) and the fisheries are not economical except when the stock is large. The TAC has therefore not been caught and shrimp can not be transferred to other species. What causes high cost of fuel in shrimp fisheries is relatively even distribution, difficult to locate denser aggregation by acoustics and small mesh in the gear.

Some species affect bycatch by scratching the fish they are caught with, good example is redfish. This is less of a problem in freezing vessels where the catch is skinned before it is frozen. Value of cod and haddock that is landed fresh but skinned decreases on the other hand considerably. For cod fisheries the problem can be solved by increasing mesh size when catching in areas where cod and redfish is to be expected. Use of large mesh requires high proportion of large cod, else more goes through the meshes than is caught which is a situation one wants to avoid. The goal of high proportion of large cod in catches is best reached by low harvest rate. Another side of the problem appearing recently is that most of the redfish (*Sebastez norwegius*) is becoming very large (recruitment failure, old cohorts last well) making separation from cod difficult.

Bycatch in deepwater shrimp fisheries (Greenland halibut and cod) has to be separated from the shrimp catch and put in a separate cod end, else the value of the catch is severely reduced. This separation was first allowed few years ago but by-catch is necessary for the shrimp fisheries to be profitable. It can be discussed if by-catch should be used to justify shrimp catch but large part of fisheries are mixed fisheries where total value of the catch matters. Cod as by-catch in gillnets that lie for a long time (days), like nets to catch lumpsucker, Greenland halibut and monkfish has low value, but today nets targeting cod lie only for few hours to have the fish alive when gutted.

Among the factors that make catches of certain species economical are:

- Tendency for forming schools, especially if they can be located by acoustics or the location is known apriori. Good examples are cod and redfish. The schools can disappear if fishing pressure is high.

- Species with relatively even distribution and/or species that can not be located by acoustics are expensive to catch. Example is deepwater shrimp where fuel cost is high and the fisheries are quickly unprofitable.

- More expensive is to catch small animals than large (except the density is extremely high), smaller meshes lead to more resistance in the gear. Example deep water shrimp.

- Species caught by many different gears are easier to catch (more difficult to avoid). Cod is caught by nearly all demersal gear used in Iceland and haddock are the most important one. What types of gear are used in each area depends of course on the most important species.

- Stock size is an important factor and what is called “target species” is usually rather abundant one. Still, captains could be searching for a species that is not a large part of the catch but they are still trying to maximise the catch. Stock size and catch in this context refers to values.

- Species that are expensive are more likely to be overfished, more so if they are easily caught.

Looking at cod in this context, it is a valuable species with tendency for schooling and is caught by many types of gear.

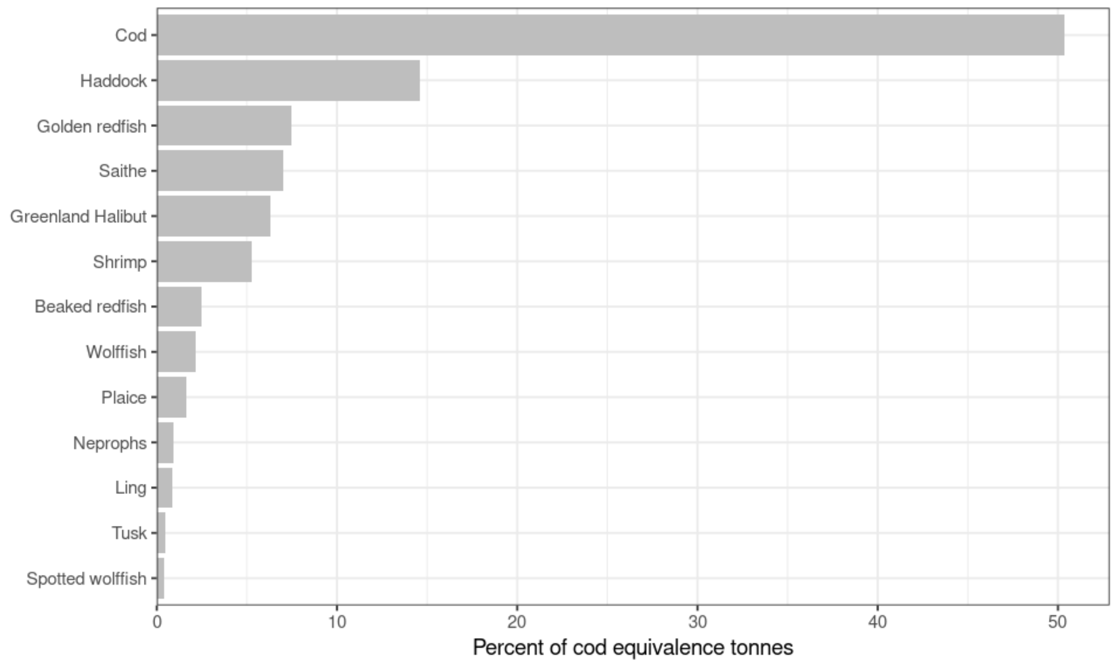

Figure 10.2: Distribution of cod-equivalence tonnes of demersal fish in Iceland 1994-2019.

Looking at the period 1994-2019 cod accounts for 50.3% of total value (cod-equivalence tons) from the demersal fleet, haddock 14.6% , golden redfish 7.5%, saithe 7.0%, Greenland halibut 6.3%, shrimp 5.2% and nephrops 0.9% (Figure 10.2). The variability in those percentages has been substantial during this period, before 2000 shrimp accounted for 13% of the value. Cod has been 55% of the value since 2010 and more than 60% last 5 years. Haddock was 20% of the total value from 2002-2011 or ~20% but catch of cod was low in this period (Figure 10.3)

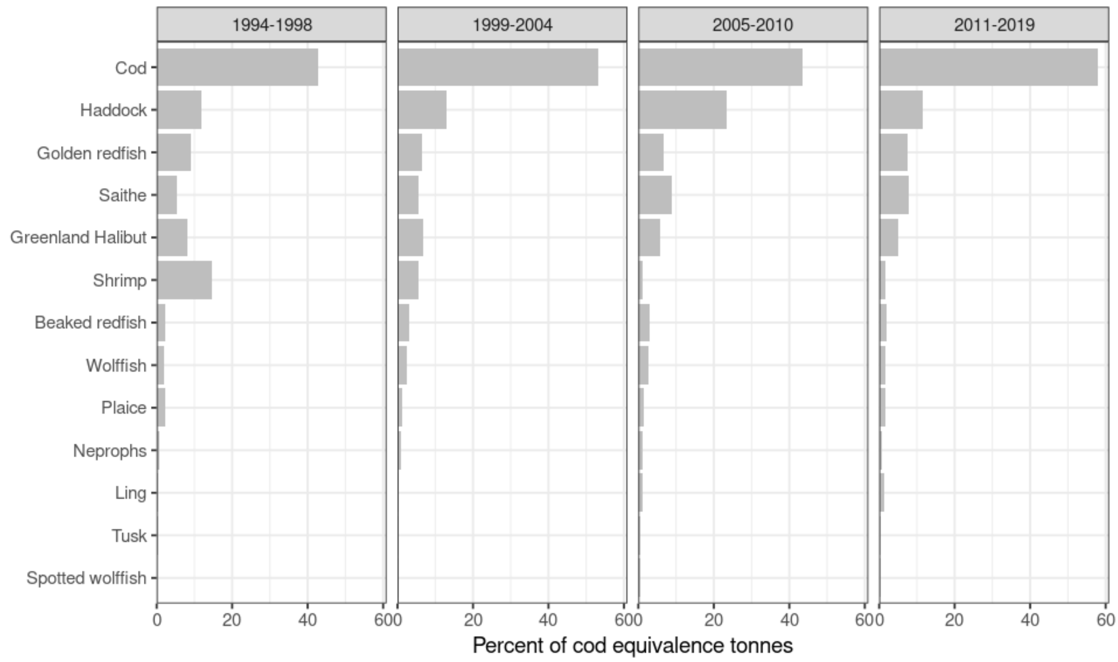

Figure 10.3: Distribution of cod-equivalence tonnes of demersal fish in Iceland in different periods.

Looking at the period 1994-2019, average catch of cod was 211 thous. tonnes and relative standard deviation 16%. Combined catch of cod and haddock combined were on the average 272 thous. tonnes with relative standard deviation of 10%. In this period the time of highest catch of haddock was the time of lowest catch of cod so the relative standard deviation of combined catch is low.

Share of cod in total value would probably increase if cost of the fisheries was included in the "cod equivalence factor." Then the value of cod compared to other species would increase and transfer from other species towards cod which is not allowed now would be possible. It must though be mentioned that finding a "cod equivalence factor" that takes into account cost of fisheries is not simple, some kind of exchange market for a part of the total TAC for each species could be a possibility.

Logbooks are here used to try to find out which species are really bycatch in cod fisheries. First, all records for demersal gear since 2010 are used, but as mentioned above the share of cod was high in this period. Sometime the word catch is used below but figures 10.4 to 10.6 are all in cod equivalences values, does not matter when comparing cod and haddock that have nearly the same cod equivalence coefficient.

Figure 10.4 shows how much bycatch different species are. The results are based on logbooks since 2010 where demersal gear are applied. Looking at Greenland halibut ~20% of the catch is caught in records where less than 80% of the total catch (in codequivalence tons) is Greenland halibut (> 80% in records where 80% of the catch is Greenland halibut). Comparable ratio for cod is 30%. Cod and golden redfish are according to this figure the species mostly caught in direct fisheries. The position of cod in this figure is though partly governed by the dominance of cod in total value of the fisheries.

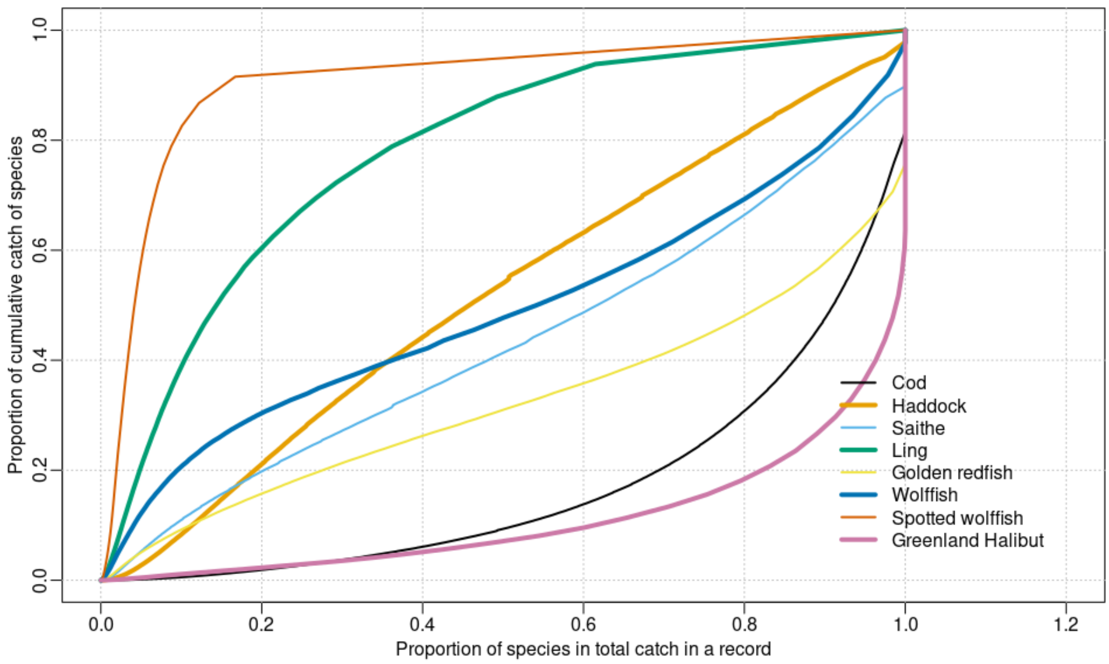

Figure 10.4: Proportion of cumulative catch of species against the proportion of the species in catch for each record. Proportion based on ration in value based on cod equivalence coefficients.

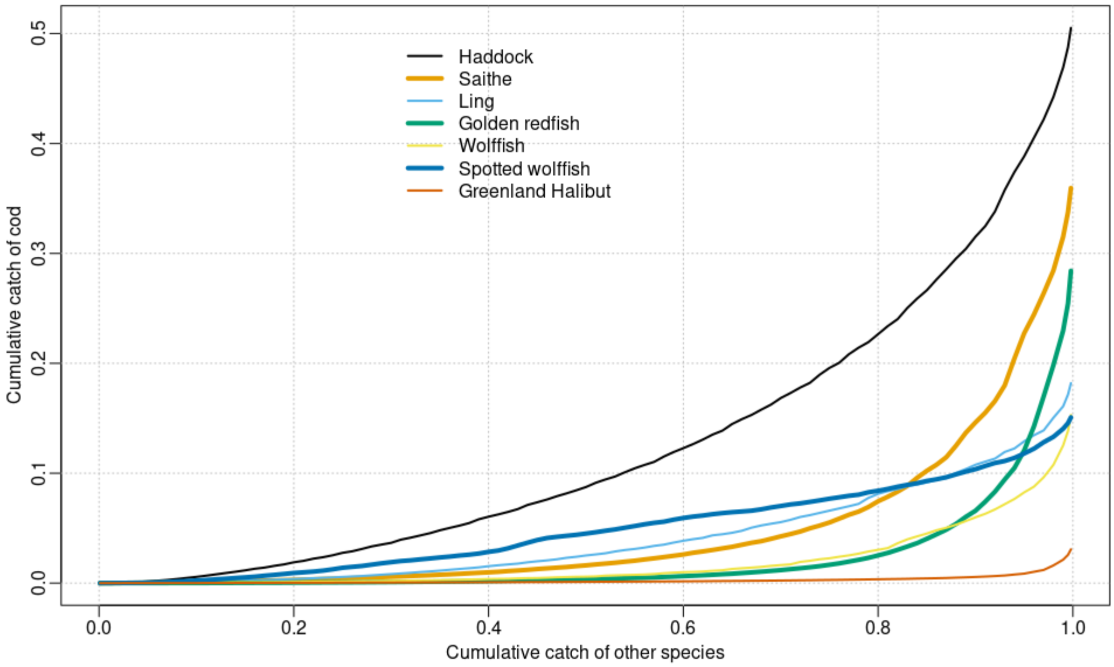

Figure 10.5: Proportion of the catch of each species caught where specified proportion of cod catch is caught.

Figure 10.5 shows how much of the catch of cod is taken with each species. 22% of the catch of cod is taken in records accounting for 80% of the catch of haddock. But the value of cod in those records in comparable to the value of haddock. Greenland halibut are the fisheries having lowest bycatch of cod and the fisheries for golden redfish are also relatively direct. In can also be seen that 50% of cod is caught where no haddock is caught.

Figure 10.6 is set up in a different way. There it can be seen that 40% of the catch of haddock and 10% of the catch of golden redfish are caught in records accounting for 80% of the catch of cod. Spotted wolffish is as much bycatch in cod fisheries as cod itself. Ling is also mostly bycatch in cod fisheries, ling is mostly caught by longliners but the catch is not enough for ling only to cover cost of bait and vessel, even though spots fishing areas with high proportion of ling could be found. Bycatch of ling is less in haddock fisheries, probably due do different depth distribution of ling and haddock.

Fisheries for Greenland halibut and golden redfish are to a large extent separated from cod fisheries. Fisheries of golden redfish have been going well in recent years, the catches have mostly been taken from dense schools that are reasonably easily located. As mentioned before redfish is avoided in cod fisheries, at least where the cod is landed fresh.

Of the demersal fish species saithe has been the most problematic species to catch. As much as is allowed of the ITQ has been transferred to other species but still large part of the TAC “has not been used.” This is as mentioned not a new problem and the fleet has developed as less of a “saithe fleet” where longlines have replaced bottom trawl and gillnets. Price is also relatively low, cod bycatch not wanted in the saithe fisheries and the stock size could be overestimated.

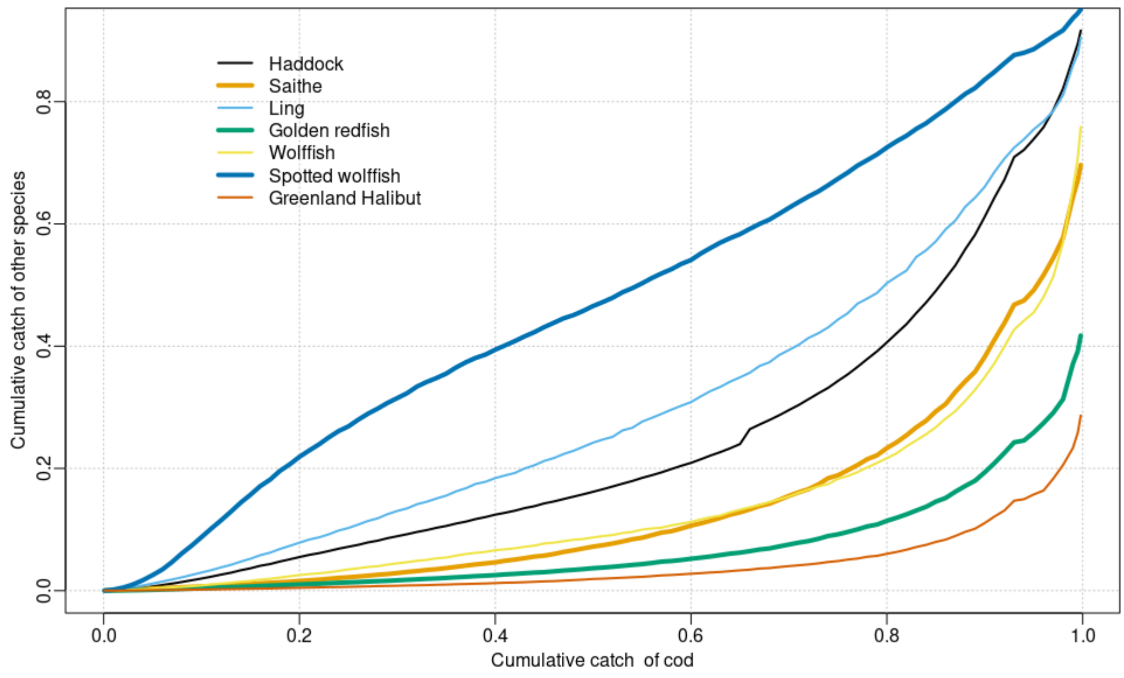

Figure 10.6: Cumulative proportion of the catch of cod vs cumulative catch of other species.

Analysis shows that cod is by far the most important species in demersal fisheries and most species other than Greenland halibut, shrimp, *nephrops*, beaked redfish, golden redfish and perhaps saithe are just part of cod fisheries.

Fisheries for Greenland halibut have usually been trawl fisheries in very deep waters where the main bycatch has been beaked redfish but Greenland halibut has also been caught as bycatch in offshore shrimp fisheries. Some cod is caught as bycatch in Greenland halibut fisheries and the amount is considerable even though it is low proportion of cod catch. The data could be investigated with regard to depth to see if the main target species is cod or Greenland halibut, the answer might just be that the captains are looking for mixture. Price of Greenland halibut is high (high cod equivalence coefficient Figure 10.1) so profitable fisheries can be conducted even though CPUE is relatively low.

In recent years increasing proportion of Greenland halibut has been caught in gillnets. According to logbooks Greenland halibut is more than 99% of the catch in those fisheries but comparable proportion in trawl fisheries was 86%.

To summarize, fisheries for most demersal species in Iceland can be seen as bycatch in cod fisheries and the harvest rate for those species has to be adapted to the cod fisheries. An exception is Greenland halibut but fishing areas are separate from other species and the same does also apply to beaked redfish. In last 15 years the fisheries for golden redfish have also been largely separated from cod fisheries and high proportion of the catch taken from dense aggregations. Their existence is probably dependent on moderate fishing pressure and redfish is as mentioned before not popular bycatch. Saithe is a problematic species, fishing the TAC does not seem to pay off.

10.3 Cost of fisheries.

Data from Statistics Iceland from the years 1997-2018 on income and cost of fisheries were used. They are split up in categories (vessels less than 10 grt, vessels 10-200 grt, vessels > 200 grt, pelagic fisheries, fresh-fish trawlers and freezing trawlers). For each category and year, the factors: Total income, fuel cost and cost of fishing gear are given.

Pelagic fisheries are not included in the following analysis but *nephrops* and shrimp vessels and some other fisheries nearly completely separated from cod fisheries are included (they are not separated in the data) (figures 10.7 and 10.8). The cost of fishing gear is more stable than the cost of fuel and has as proportion decreased with time. Variability in fuel cost as proportion of income can reflect variability in price of oil, amount of oil used or price of products.

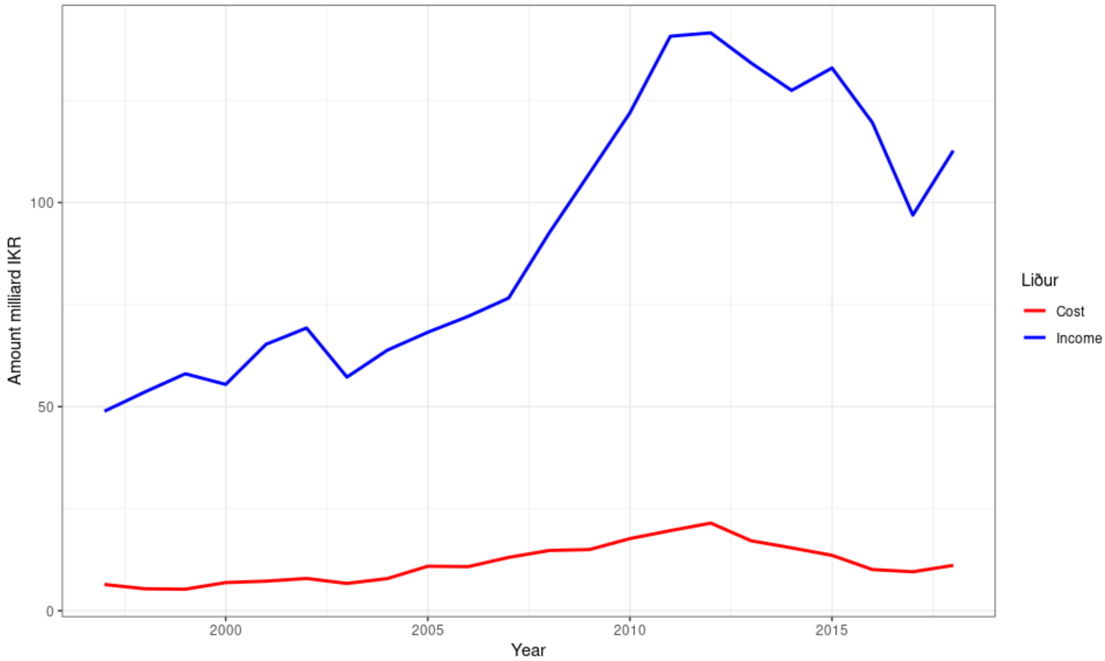

Figure 10.7: Income and cost of fuel and gear for the Icelandic demersal fleet 1997-2018.

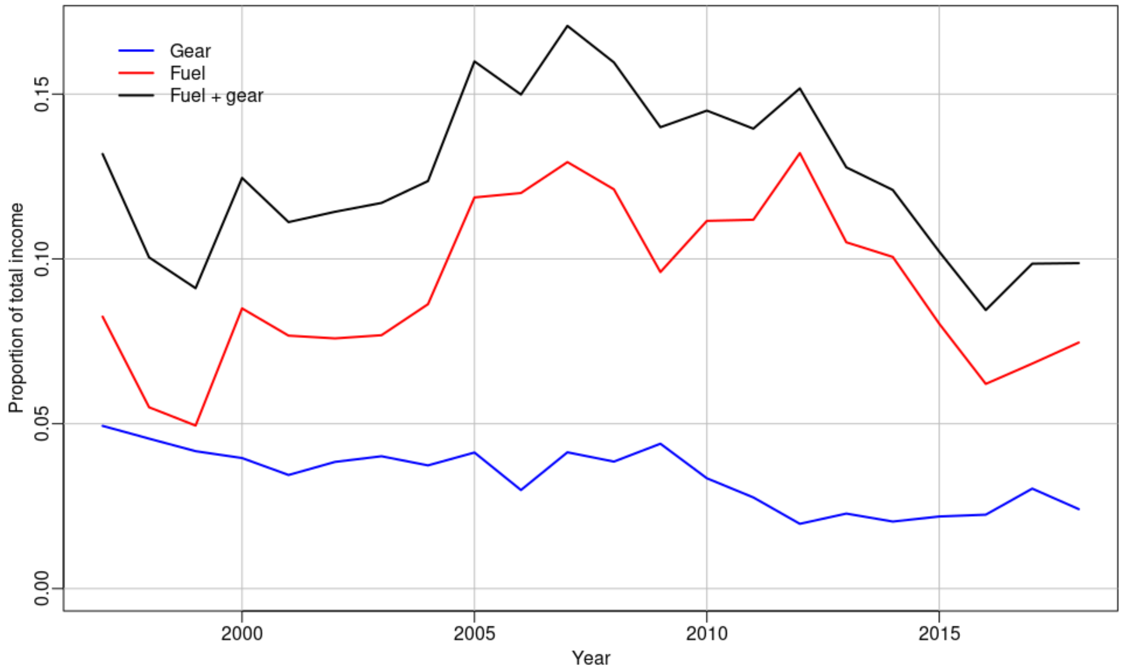

Figure 10.8: Fuel and gear cost of the Icelandic bottom trawl fleet as a proportion of total income.

CO_2 relased by Icelandic fishing vessels is shown in Figure 10.9. Pelagic vessels are included in those numbers and have substantial weight is some years, especially after fisheries for blue whiting increased. Fishers by Icelandic vessels in the Barents Sea are also included in those numbers an also numbers on fuel cost and total income for fishing vessels.

From 1991-1997, 30-55 thous hours were annually trawled in fisheries targeting Greenland halibut, usually large conducted by large trawlers using large trawl in deep waters. In addition, 15. thous hours were trawled annually in the Loop hole in the Barents Sea from 1994-1997 compared to 5000 hours in the Barents Sea last 15 years. From 1993-1995 300 thous. hours were trawled annually targeting offshore shrimp (on the average small trawlers) and more than 200. thous from 1991-2000. Those factors explain much CO_2 release by the fishing fleet from 1993-2000.

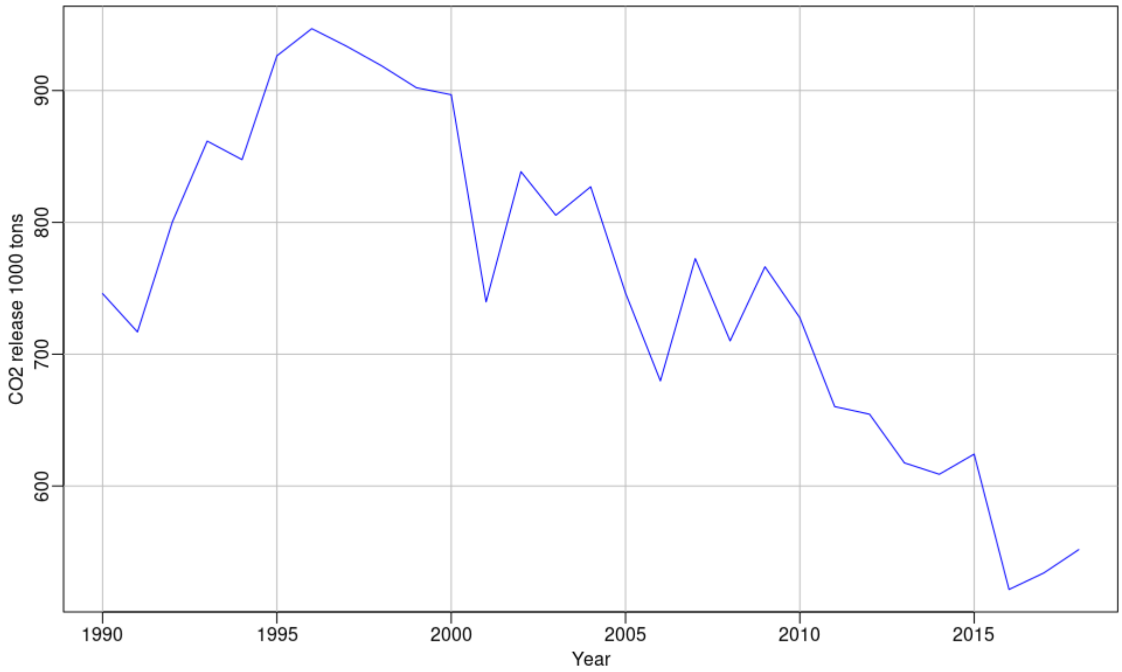

Figure 10.9: CO2 release by Icelandic fishing vessels 1990-2018.

Price of fuel to fishing vessels can be calculated by dividing the CO_2 release of the Icelandic fleet by the cost of fuel obtained from Statistics Iceland. Finally, the result is multiplied by 2.64 to take into account that 1000 liters of oil are equivalent to 2.64 tonnes of CO_2 (Figure 10.10).

The number obtained this way is quite variable. Part of the variability is related to the Icelandic krona where major drop in the value occurred between 2008 and 2009. Converting the results to SDR or some other currency could be an option (Euro did not exist before 1999). But the main problem is that the price of fuel obtained this way seems too low, whatever is the reason. Cost of fuel (~10% of income Figure 10.10) seems relatively low although it matches well with the numbers obtained from the analysis done here basing fuel consumption per hour trawled on (Björnsson (2004)). This discrepancy needs to be investigated, and part of the reason for using the data on CO_2 release and cost from Statistics Iceland was to check consistency of different data sources. If the use of fuel is higher than predicted here, optimum Harvest rate would decrease.

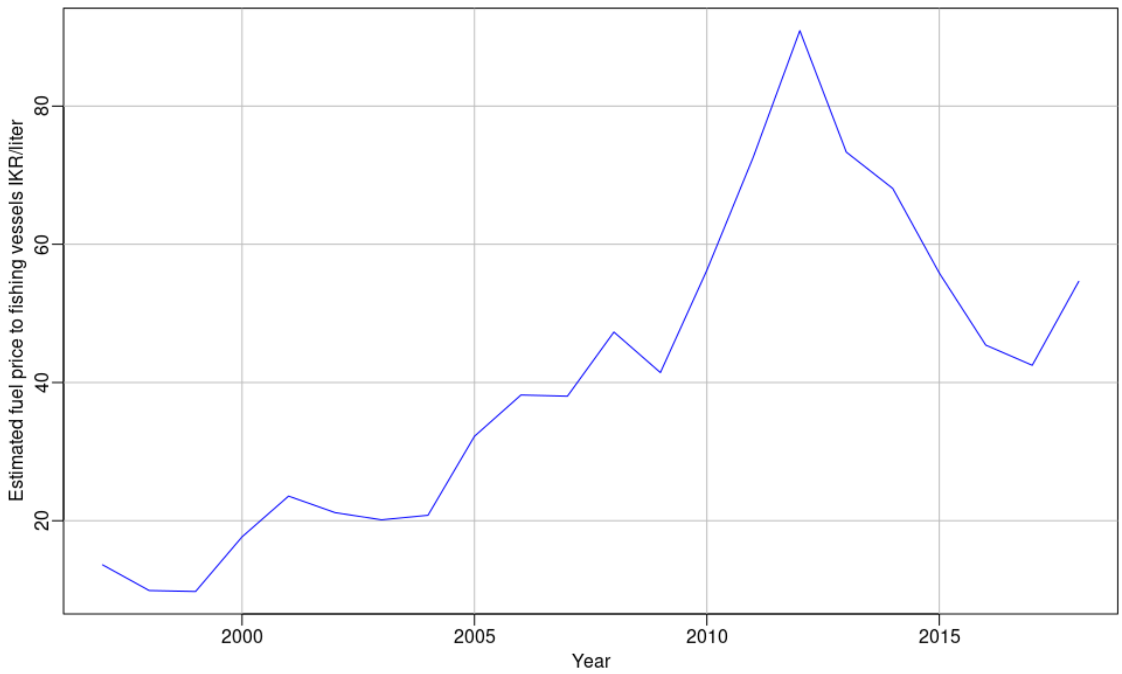

Figure 10.10: Estimated fuel price to fishing vessels IKR/liter.

10.4 Estimation of maximum revenue.

In the section above cost of fuel and gear as proportion of total income was estimated for the years 1997-2018. The result is 8-17% of income, relatively low cost and estimated price of fuel is rather low (Figure 10.10). Saving of more than 5% of total income can therefore not be obtained by reducing the cost of fuel and gear. The relationship between stock size/harvest rate and CPUE is the decisive factor when estimating cost reduction/increase and this relationship is also key equation in stock assessment.

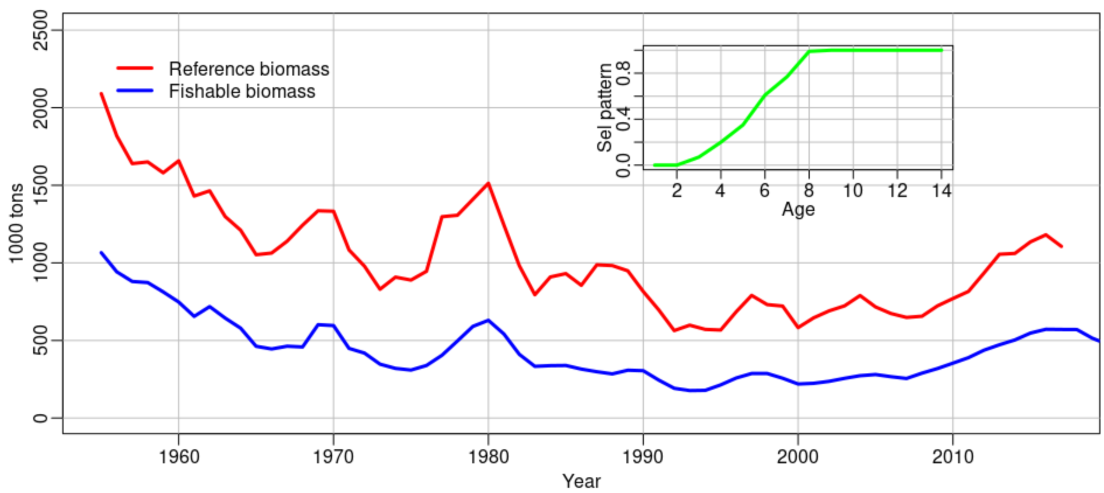

Figure 10.11: Development of fishable biomass and reference biomass 1955-2017. The selection pattern used to compile the fishable biomass is shown in the small figure.

In the evaluations of HCR for Icelandic cod 1994 and 2004 cost per unit effort was assumed to be fixed but catch per unit effort increased by the fishable biomass in the power 0.7, the increase was estimated from CPUE data from the fleet account for landing trips. Fishable biomass is defined in the traditional way i.e from average number of fish in the year, estimated selection pattern of the fleet and mean weight at age in the catches. $Fb_y = \sum_a N_{y,a} \frac{F_{ya}}{Z_{ya}} Sel_a cW_{y,a}(1 - e^{Z}_{ya})$ where Sel_a is average selection pattern of the fleet. The fishable biomass is considerably smaller than the reference biomass (B_{4+}), closer to the spawning stock.

Results from surveys indicate that the power in the relationship between catch of cod and size of fishable stock could be higher than 0.7 or $1.35 - 1.45$ (10.12). This mean that 50% larger fishable stock can lead to $1.5^{1.4} = 1.76$ or 76% higher CPUE. Catch rate could increase even more then in the surveys as when fishing pressure is low cod (and some other species) can form dense schools that captains can locate with acoustic equipment. Higher catch rate is though not fully reflected in average catch per day, the catch per haul will be similar so the haul will be shorter (too much catch per haul reduces quality), capacity in gutting and handling of fish is limited, storage capacity in the vessel is limited and the catch rate of other species is not as high. Taking notice of those factors catch per day is assumed to be proportional to the fishable stock.

As CPUE does always increase with larger stock but the cost of unit effort is identical, the harvest rate leading to maximum revenue (HR_{mse}) is always lower than the harvest rate leading to maximum catch (HR_{msy}). How different they are depending on a number of factors.
* Higher power in relationship between CPUE and fishable stock leads to lower HR_{mse}.
* Higher cost as proportion of income leads to lower F_{mse}.
* Linking price of fish to size or stability in supply leads to lower HR_{mse} as lower HR leads to larger fish and less variability in catches.

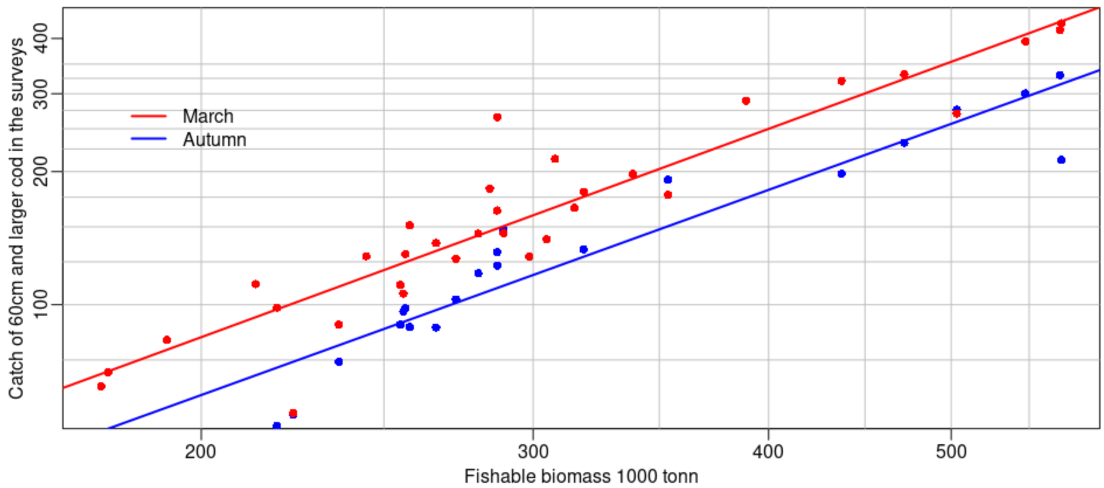

Figure 10.12: Catch rate in surveys kg per hour against estimated fishable stock. The lines shown that are fitted to all data until 2017 have a slope of 1.35 for the March survey and 1.45 for the autumn survey. The slope of a line fitted on log scale becomes power on normal scale to CPUE in the surveys is proportional to fishable biomass in the power 1.35-1.45.

Premises about price can be based on data from Directorate of fresh fish prices or fish markets. According to data from Directorate of fresh fish prices from January 2015 (Figure 10.13) the price of ungutted cod varied from 155 IKR/kg for 1.0 kg cod to 250 IKR/kg for 3.7 kg cod 300 IKR per kg for 6.5 kg and larger or $154.87 + 38.1 \times (W - 1)$ if $W < 3.5kg$ and $250.18 + 16.7 \times (W - 3.5)$ if W is between $3.5 - 6.5kg$. The equation used 2003-2004 and data from Directorate of fresh fish prices from 2019 show similar development in price with size of cod. The price for larger cod is always higher.

Following the work of earlier HCR evaluations, wages of the crew are not considered cost in those analysisis, if that was done the harvest rate leading to maximum revenue would be lower as the wages are specified proportion of the price of the fish. The wages of the crew are cost from the viewpoint of the vessel owner so the lower HR_{mse} obtained this way would the the harvest rate maximising the MSE of the vessel owners.

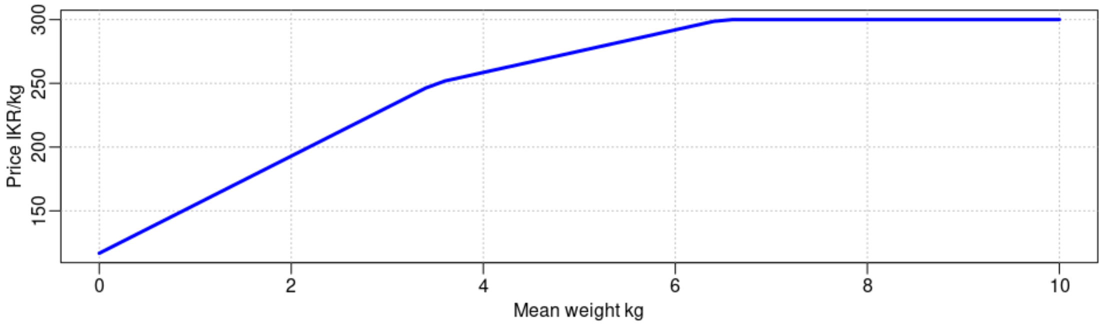

Figure 10.13: Price of cod in January 2015 according to data from Directorate of fresh fish prices.

Use of oil is a factor that will be monitored in the future regarding impact of fisheries on the environment and fuel cost is an important part of total cost of the fisheries. In a thesis by Eyþór Björnsson at University of Akureyri (Björnsson (2004)) and various lectures and articles by his supervisor Emil Ragnarsson, fuel consumption of fishing vessels is evaluated, for example on how it is split between gear and the vessel and also between different parts of the fishing gear. According to (Björnsson (2004)) use of oil in fisheries by demersal trawl was 0.50l/kg 1998-2002, increasing from 0.44 l/kg 1994-1996. The increased in this period was explained by increasing proportion of freezing vessels but freezing of catch requires considerable energy. Based on oil price 100kr/l the cost of oil would have been 45kg/kg (at current value) in the year 2000. As may be seen in Figure 10.10) calculated price of oil has often been lower than 100kr/l but investigation on the price of diesel fuel for vessels confirms that 100 kr/l are approximately correct. The cost of fuel must though matter, some of the newest fresh fish trawlers have propellers close to 5m with a engine power of only 1800 kw, the large propeller only put in to reduce fuel consumption.

She share of longliners i demersal fisheries has increased since 1996. Since 2010 35% of the catch of cod has been taken by longliners, 43% in bottom trawl, 9% in gillnets, 7% by handliners and 6% in demersal seine. The use of fuel in longlining is less than for trawlers or 0.15-0.2 l/kg in the year 2000 (Björnsson (2004)). Cost of bait is though considerable. Assuming 10 kg of bait and 100 kg caught per 500 hooks (the unit tub) and 200 kr/kg price of bait, the cost of bait is 20 kr/kg. 100 kg/500 hooks is relatively low catch but realistic when the stock is small.

The result of those analysis is that the cost of fuel and bait for longliners is comparable as cost of fuel for trawlers. Cost of bait can be expected to decrease when harvest rate is increased as catch per hook increases, both due to higher number of fish and higher mean weight.

When calculating the cost of fuel, gear and bait the base number for the year 2000 is 55kr/kg (0.55 liters/kg) and that the cost is inversely proportional to the fishable biomass to the power 1.0.

Cost of aquiring and maintaining vessels is large part to total cost of fisheries and in Iceland relatively low proportion of this cost is domestic. Many companies operate relatively old vessels that have been well maitained or even reconstructed. For a company operating specified number of vessels depreciation and maintainance can be looked at as fixed cost and saving in those factors can only be obtained by reducing the number of vessels. To get some idea about the cost of vessels the following example was set up.

- Cost of acquiring a fresh fish trawler 2,5 milliard IKR, based on the price of 3 trawlers built for the company HB-Grandi in 2015.

- Depreciation + maintenance 10%. Interest 3% of original price, more when the vessel is new , less with older vessel. Annual cost per vessel is according to those assumptions 320 million IKR.

- Annual catch of demersal fish for each vessel is 6000 tonnes based on the years 2012-2014. This a high value as only 11 trawlers caught this amount 2012-2014. It is assumed that cod is 50% of the value of the catch.

Based on the premises above cost of vessel per kg caught is $\frac{320}{6} = 53$ IKR based on the years 2012-2014. This cost is estimated to be inversely proportional to the fishable biomass to the power 0.8. This cost is based on trawler but assumed in the long run to be similar for other type of vessels.

The equations included in the model were.

- Cost of oil and gear $C_1 = 55 \times \frac{B_{c,2000}}{B_c}$
- Vessel cost $C_2 = 53 \times (\frac{B_{c,2012-2014}}{B_c})^{0.8}$
- Price per kg (P_w) based of values from Directorate of fresh fish prices from January 2015 (figure @ref(fig:pricejan2015)
- Income $I = C \times P_w$
- Revenue $H = I - C \times (C_1 + C_2)$.

Where B_c is fishable stock, $B_{c,2000}$ fishable stock 2000 (228 kt) , $B_{c,2012-2014}$ average fishable stock 2012-2014 (450 kt), C_1 is cost of gear and oil per kg fish, C_2 cost of vessel purchase and maintainance per kg fish, P_w price per kg as function of mean weight W, C is catch, I value of catch and H revenue. The case where price per kg is fixed are also investigated, both 250 kr/kg and 170 kr/kg (Table 10.1).

The result is that based on price from January 2015 maximum value of catch is obtained with Harvest rate of 0.19. Maximum revenue excluding vessel cost is obtained with HR = 0.18 but 0.15 if cost if vessel cost is included. Fixed price of 250 IKR/kg gives maximum value when the catch is at maximum i.e with harvest rate of 0.26 (Table 10.1). Maximum revenue with fixed price of 250 IKR/kg is reached at 0.19 only gear and oil cost is included but at 0.17 if vessel cost is included. Lower price leads to maximum revenue obtained at lower Harvest Rate. In earlier chapters, maximum catch is reached at lower harvest rate or 0.23 but those results are based on $B_{trigger} = 0$ but the results here on $B_{trigger} = 220$ thous. tonnes.

The curve closest to be called national revenue is the green line "Value of catch - oil, gear and vessel" giving maximum at HR = 0.15. Wages of the crew are also included in cost for the vessel owners and maximum revenue is in that cast reached at HR lower than 0.15. (red curve).

Looking at the curve including all cost of the vessel owners (red curve in Figure 10.15) the revenue is 50% of value of catch with HR in the range 0.14-0.20 and national revenue exceeds 70%. Value of catch and therefore wages of the crew change only by 15% in the same range, mostly because of higher price for larger fish.

With $B_{trigger} = 0$ the vessel the vessel owners start losing money if HR exceeds 0.34 but at higher harvest rate with $B_{trigger} = 220$ as the realized harvest rate will always be lower than intended as the spawning stock is usually below $B_{trigger}$

Taking cost of the fisheries into account does lead to revenue at HR below 0.2, even though only cost of oil and gear is included.

Variability in national revenue caused by variability in stock size is more than the difference between average values (Figure 10.16). The variability does always decrease as HR is reduced. The figure does not include variability in price of products and supplies nor how the size of the fishing fleets follows variability in stock size. In principle similar number of vessels is required independent of the level of the stock as long as the HR is the same.

Investigation of data from the Statistics Iceland (figures 10.7 and 10.8) show that the cost of oil and fishing gear was on the average 12.5% of the value of catch from 1997-2018. According to the current analysis will this part of the cost be 12.5% of the value of catch if HR is 0.25 (Average HR 1997-2018) similar to the value obtained using the data from Statistics Iceland. The analysis done here indicate that this ratio would be 8.2% if HR is 0.2.

Similar analysis were conducted in 2015. Premises about future recruitment have not changed much since 2015 and the same data on cost and price of fish are used, so the results now are similar.

The cost of fishing is probably underestimated in the analysis here (the revenue is often very high proportion of the income) but more cost would further decrease the HR giving maximum revenue. The size of the cod stock will increase considerably if the HR is decreased to 0.18 and lower so factors like density dependent growth and predation on cod and other species would have to be investigated.

The analysis done here are thought of as indicators on changes in national revenue when effort is reduced and the fleet reduced accordingly. The premises are that large fish stocks allow to get the same amount of fish and more value with less cost , other than the wages of fishermen that in Iceland are proportion of the value of catch. The analysis done here are very simple and only intended to give idea about optimum HR, but not about future revenue of the fisheries. If cost increases or price of fish decrease from premises used here, revenue will decrease as will the harvest rate giving maximum revenue.

The factor that has largest effect on most economical HR is relationship between CPUE and stock size, also key equation when estimating size of fishstocks. In addition, some coherence in HR of different species is assumed so mixed fisheries problems will not lead to reduction in the economics of the fisheries.

The main conclusion from those investigations is that the harvest rate in current management plan is somewhat higher than the harvest rate giving maximum revenue.

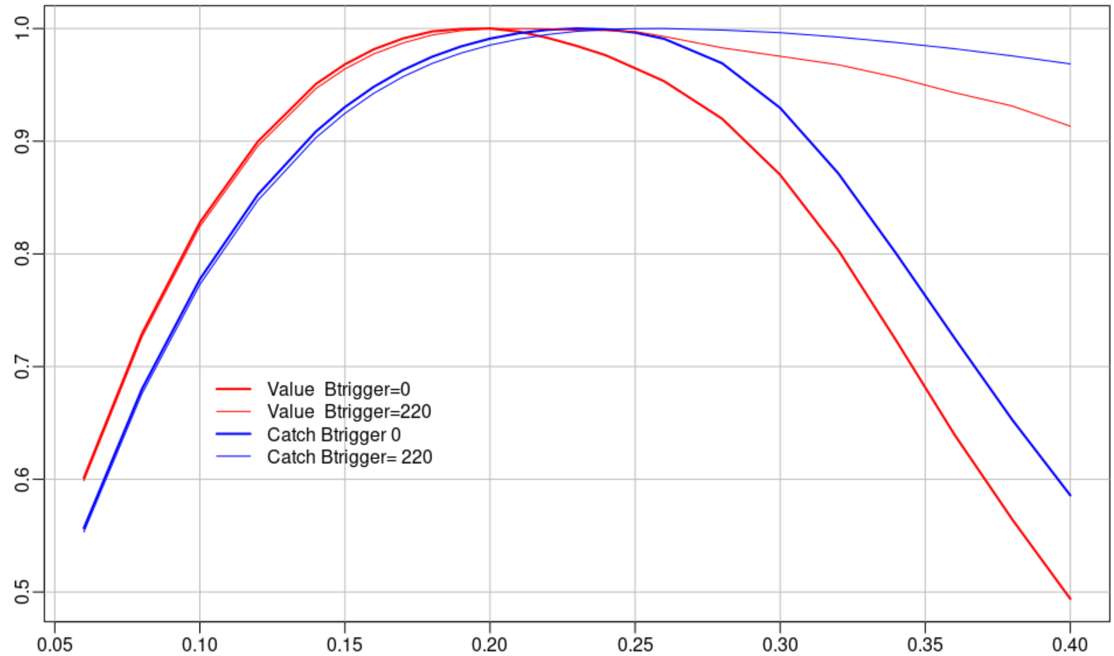

Figure 10.14: Value of catch against Harvest Rate, all the curves are scaled to a maximum of 1.

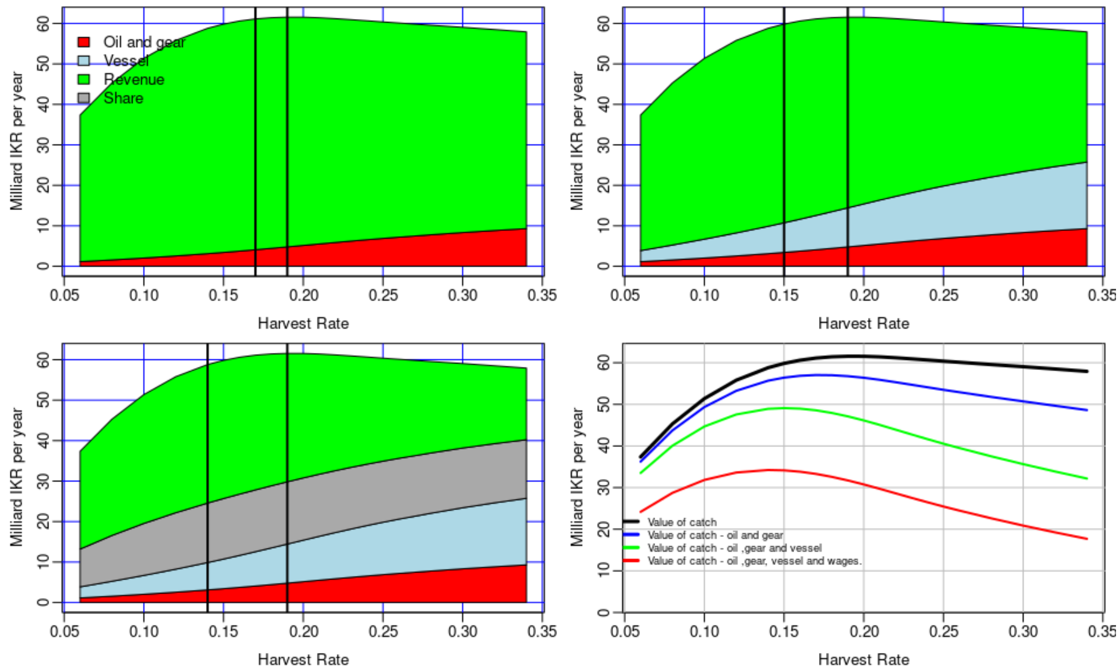

Figure 10.15: Cost, total income and revenue against harvest rate based on price per kg from January 2015. The vertical linse show the harvest rate giving maximum value and maximum revenue (green area). Btrigger=220 kt

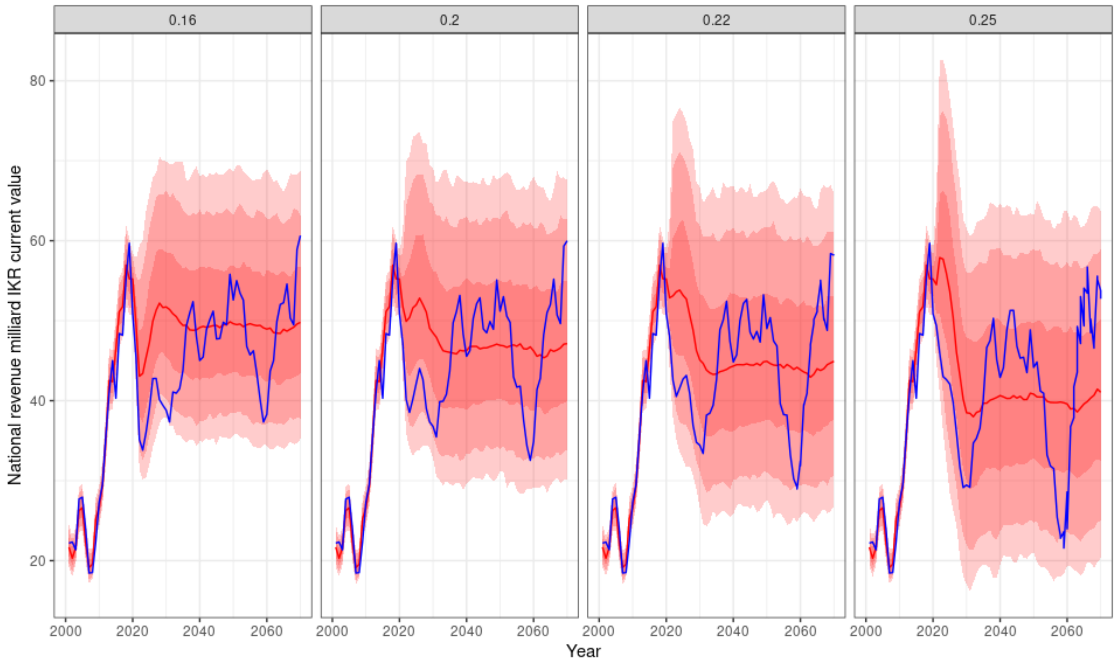

Figure 10.16: National revenue i.e. value of catch minus cost of oil, gear and vessel base on 4 different harvest rates. Variability in price of fish and oil not taken into account. Average of future values are shown as thick red line but the shaded areas show 50,80 and 90% of values. The blue line show one future replica.

Table 10.1: Harvest rate giving maximum revenue based on price per kg from January 2015, fixed price of 250 IKR/kg and 170 IKR/kg. Values including vessel cost and excluding it are shown.

Oil and gear	Vessel	Price	Max income	Max revenue
x		Variable	0.19	0.17
x	x	Variable	0.19	0.15
x		250	0.26	0.19
x	x	250	0.26	0.16
x		170	0.26	0.18
x	x	170	0.26	0.14