BENCHMARK WORKSHOP ON BALTIC COD STOCKS (WKBALTCOD2)

VOLUME 1 | ISSUE 9

ICES SCIENTIFIC REPORTS

RAPPORTS SCIENTIFIQUES DU CIEM

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

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

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

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

ICES Scientific Reports

Volume 1 | Issue 9

BENCHMARK WORKSHOP ON BALTIC COD STOCKS (WKBALTCOD2)

Recommended format for purpose of citation:

ICES. 2019. Benchmark Workshop on Baltic Cod Stocks (WKBALTCOD2).
ICES Scientific Reports. 1:9. 310 pp. http://doi.org/10.17895/ices.pub.4984

Editors

Michele Casini • Meaghan Bryan

Authors

Alessandro Orio • Anastasiia Karpushevskaia • Anders Nielsen • Andreas Sundelöf • Casper Willestofte Berg • Christoffer Moesgaard Albertsen • Conrad Stralka • Francesca Vitale • Franziska Schade • Fritz W. Köster • Hans Jakob Olesen • Harry Strehlow • Hege Sande • Henrik Mosegaard • Jan Horbowy • Jane Behrens • Joakim Hjelm • Johan Lövgren • Jonna Tomkiewikz • Karin Hüssy • Kate McQueen • Keith Brander • Kåre Nolde Nielsen • Marc Simon Weltersbach • Marcin Rucinski • Margit Eero • Maria Pierce • Marie Plambech • Marie Storr-Paulsen • Maris Plikshs • Massimiliano Cardinale • Michael Andersen • Michele Casini • Mikaela Bergenius • Monica Mion • Nils Höglund • Sebastian Linke • Sofia Carlshamre • Staffan Larsson • Stefanie Haase • Uwe Krumme • Verena Trenkel • Vladlena Gertseva • Viktoriia Amosova • Xochitl Cormon • Yvette Heimbr

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

Contents

i Executive summary

The ICES Benchmark Workshop on Baltic Cod Stocks (WKBALTCOD2) met at ICES Headquarters in Copenhagen, Denmark on 4–8 February 2019, following a data evaluation workshop (Chair Johan Lövgren) and several preparatory web conference meetings. The meeting, co-chaired by Meaghan Bryan, USA (External Chair) and Michele Casini, Sweden (ICES Chair), was attended by two invited external experts, Vladlena Gertseva, USA, and Verena Trenkel, France, and 47 participants from 10 countries. Participants represented a diversity of groups including industry, NGOs, managers, and scientists.

The objectives of WKBALTCOD2 were to evaluate the appropriateness of the data and the assessment methods to determine stock status for the Western Baltic cod (SD 22-24) and the Eastern Baltic cod (SD 24-32) stocks, evaluate the short-term forecasting methods, re-examine and update the reference points, and update the stock annex as appropriate to these stocks.

The workshop started with a group discussion of data issues and decisions that were integral to both assessments. Subsequent, stock-specific discussions were held in smaller subgroups and the main results and conclusions were presented in plenary. Around 15–20 persons, together with at least one external reviewer/co-chair, participated in each of the subgroups.

A single modelling approach, the state-space stock assessment model (SAM), was presented for Western Baltic cod. This model has been used previously to determine stock status of Western Baltic cod and the panel agreed that this model should be used for the current assessment. The main issues that emerged and were addressed during the workshop for this stock were stock mixing in SD 24, the inclusion of new recreational data from Sweden and Denmark, the inclusion of a German pound-net survey index of age-0 fish in the assessment model, the extension of the time-series of the assessment back in time, and the update of reference points. Future efforts should be made to update the mixing proportion of Eastern and Western Baltic cod in SD 24 by length and season, and continue improving the stock splitting methods and the geographical coverage of the samples.

An analytical quantitative assessment had been lacking since 2014 for the Eastern Baltic cod stock. The key issues addressed during the benchmark included how to best account for changes in productivity (e.g. growth, mortality, maturity) in the assessment model, stock splitting in SD 24, the use of age–length keys in the assessment, and ageing error and bias. Stock assessment models using Stock Synthesis and the stochastic surplus production model (SPiCT) were put forward for the benchmark as two possible model candidates. The panel agreed that the Stock Synthesis model assuming time-varying growth, natural mortality, and maturity to account for changes in productivity was acceptable to provide scientific advice, while the SPiCT model should be maintained as an alternative approach. The accepted model exhibited some residual patterns that were likely due to assuming that ages were precisely known. Stock Synthesis can accommodate an ageing error matrix to account for precision and bias. Ageing error and bias statistics should be developed and thoroughly reviewed for the next benchmark. Additionally, future work should focus on improving the growth estimates, which will allow a more precise separation between growth and natural mortality. The Stock Synthesis model would also benefit from information on sample size associated with length distributions of commercial catches. Estimates of fishing mortality compatible with a precautionary F_{MSY} was not attainable for this stock, therefore probabilistic forecasts with Markov chain Monte Carlo methods were proposed to be used instead.

ii Expert group information

Expert group name	Benchmark Workshop on Baltic Cod Stocks (WKBALTCOD2)
Expert group cycle	Annual
Year cycle started	2019
Reporting year in cycle	1/1
Chairs	Michele Casini, Sweden
	Meaghan Bryan, USA
Meeting venue and dates	4-8 February 2019, ICES Headquarters, Denmark (47 participants)

1 Introduction

The chairs Meaghan Bryan (USA) and Michele Casini (Sweden) welcomed the meeting participants (Annex 1).

The chairs introduced the goals and focus of the meeting and the state of the different tasks to be conducted by the group. WKBALTCOD2 has been given the following Terms of References:

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.
b) Agree and document the preferred method for evaluating stock status and (where applicable) short-term forecast and update the stock annex as appropriate. If a category 1 assessment method can not be agreed, then an alternative method (the former method, or following the ICES data-limited stock approach) should be put forward as a basis for the assessment and advice;
c) Re-examine and update (if necessary) MSY and PA reference points according to ICES guidelines (see Technical document on reference points);
d) Prioritize recommendations for future improving of the assessment methodology and data collection;
e) As part of the evaluation, conduct a 5-day data evaluation workshop. As part of the data evaluation workshop consider the quality and compiling methodology for all input data for stock assessment, including catch data.

2 Description of the Benchmark Process

A series of data workshops was conducted in preparation of the benchmark WKBALTCOD2. For the Western Baltic cod, the main goals of the workshops were to (i) prepare the recreational data and (ii) to discuss the splitting method used to differentiate between western and eastern Baltic cod from commercial catches and survey data in SD 24. For the recreational data intersessional work was conducted with one physical meeting on 17–19 October 2017 to set up a work plan followed by skype meetings on 16 November, on 18 December 2018, and on 9 January 2019. The Agenda and notes from the meeting are given in WD8. For the intersessional work on stock splitting a workshop was conducted on 3–4 July 2018. The agenda and notes from the meeting are given in WD9.

For the Eastern Baltic cod, WGBFAS has continuously worked toward addressing specific ToRs that were allocated to the group since the last benchmark in 2015. Additionally, two specific ICES workshops took place, i.e. WKBEBCA (2017) and WKIDEBCA (2018). The role of WKBEBCA was mostly to synthesize new knowledge of growth and natural mortality, which were considered the key issues for EB cod assessment. WKIDEBCA more specifically addressed the question of whether the biological base knowledge was sufficient to move on with the benchmark process, with the aim to establish a quantitative (ICES category 1) assessment for this stock. The outcome of these meetings was positive, which led to the preparation of a work plan for the following benchmark process.

A data evaluation workshop for both Western and Eastern Baltic stocks was held on 15–19 October 2018 which considered the quality and methodology to obtain all input data for stock assessment.

The chairs of WKBALTCOD2 introduced the agenda, which was shortly discussed, adjusted and finally adopted by the participants. However, a flexible agenda was adopted (Annex 2).

WKBALTCOD2 started with a group discussion of data issues and decisions that were integral to both assessments. Subsequent, stock-specific discussions were held in smaller subgroups and the main results and conclusions were presented in plenary. Around 15–20 persons, together with at least one external reviewer/co-chair, participated in each of the subgroups.

3 Western Baltic cod

3.1 Stock ID and substock structure

Cod in the Baltic Sea is assessed and managed as two separate stocks, i.e. eastern and western Baltic cod, located in ICES Subdivisions (SD) 24–32 and 22–24, respectively. There is clear evidence that eastern Baltic cod regularly occur in SD 24 (Hemmer-Hansen *et al.,* 2019). Given the apparent difference in biological parameters between the two stocks, eastern cod needs to be separated from the western stock, for stock assessment purpose (Eero *et al.,* 2014). Since the last benchmark in 2013, assessments have been conducted by stock, i.e. separating between eastern and western Baltic cod in the mixing area in SD 24. Applied stock discrimination methods on catch and survey data are described in sections 3.6.2.

In the current assessment (before this benchmark) the time-series started in 1994 since stock splitting data were not available for the entire time-series. At that time, only data from the Danish commercial catch were available to conduct stock splitting.

3.2 Issue list

The main issue with the assessment of cod in SD 22-24 was to investigate if a new survey index with the east / west split included in SD 24 could improve the assessment. In the current assessment survey information east of 13 degrees have not been included in the indexes as between 13 and 15 degrees East a small proportion of WB cod is present. Since 2013, German recreational catches have been included in the assessment. Sampling of the Danish and Swedish recreational fishery has been ongoing and another large issue was to test the inclusion of these data into the assessment. A third larger issue was to extend the catch matrix back in time to ensure the coverage of a longer time-series in order to support the estimation of reference points for the stock. The complete issue list is included in WD12.

3.3 Fisheries data, multispecies and mixed fisheries issues

During the data compilation workshop the fishery of the Western Baltic cod stock was discussed.

The western Baltic cod stock has experienced large fluctuations in stock development and landings over time. In the mid-1980s, landings were close to 50 000 t in the western Baltic management area decreasing to below 6000 t in 2017. Unlike the eastern Baltic cod, there is no documentation of decreased condition or impairment by reduced growth of western Baltic cod. The western Baltic cod has experienced high fishing pressure and shown poor recruitment for several years and was assessed to be well below reference points at the onset of the landing obligation in 2015 (Valentinsson *et al.,* 2019). Although the spawning stock was at a historically low level in 2016, a new large year class was observed, which is likely to influence the development of the stock in the following years and was the reason for increased quotas in 2019. Discards in the western Baltic management areas are estimated from observer programs in Denmark, Sweden, Poland and Germany. For Western Baltic cod discard data have been included in the assessment since 2002. For a long period of time the discards of cod in the Baltic were considered relatively low compared with other areas. The average discard rate in the western Baltic cod stock was 8% (for the period 1994–2017), and has for the last three years (2015–2017) been estimated to be below 5% (ICES 2018).

Trawls and gillnets take the main part of western Baltic cod. The main fishing nations are Denmark and Germany, with Poland and Sweden as the third and fourth most important fishing nations (Figure 3.6). The majority of the landings in SD 22-23 are taken in Q1, however this pattern has changed in recent years as a spawning closure has been in place for 6–8 weeks in Q1. In an STECF report from 2017 it has been calculated that fishing effort (kw-days) of otter trawls which has substantially declined in recent years, while the effort development for gillnetters was more variable (STECF 2017)(Figure 3.7).

Western Baltic cod is mainly fished by 10–18 metre vessels, however in certain areas closer to shore also by vessels below 10 metres, this is mainly in SD 22 and SD 23 whereas in SD 22 and especially in SD 24 larger vessels are participating in the fishery (Figure 3.8). In some years where the quota has been less restrictive larger more mobile vessels from adjacent areas were participating in the fishery, however in later periods with more restrictive quotas the fishery is more dominated by local less mobile vessels.

The main part of western Baltic cod fished by a flag country is landed in the same country (Figure 3.9). This indicates that the sampling of foreign landing is not of major concern.

Recreational cod catches are mainly taken by private and charter boats and to a smaller degree by land-based fishing methods. Rod-and-line fishing with artificial lures or live bait is the primary fishing method targeting cod (Weltersbach *et al.*, 2019). Cod angling takes place throughout the year in Baltic waters. A minority of catches are taken by recreational passive gear fishers.

Regulation

The regulation has changed over time in both the Eastern and Western Baltic management area. The main changes have been for gears, where different minimum mesh sizes were introduced and minimum landing size was decreased in 2015 from 38 to 35 cm. The latter is thought to have had an effect on the discard ratio, but mainly in the eastern Baltic (Figure 3.10). Further, a spawning closure has been in place in the western Baltic area, which was changed in timing and duration over time. In 2019, the spawning closure was suspended in response to the increased stock size, resulting from the strong 2016-year class.

There are no seasonal or spatial closures regulating the marine recreational fishery for cod. The legal minimum landing size (MLS) for cod varies between countries and federal states and is 35 cm respectively 38 cm. In 2017, a bag limit was introduced limiting recreational cod harvest to three cod per day/angler during the closed season (February to March) and five cod per day/angler for the rest of the year. In 2019, the bag limit was raised to seven cod per day/angler for the entire year.

3.4 Ecosystem drivers

Hydrodynamic conditions within the western Baltic Sea are extremely variable, particularly in the narrow Belt Sea, the Sound, and the Fehmarn Belt, through which all water passes in and out of the Baltic Sea (Matthäus and Franck, 1992; Schinke and Matthäus, 1998). The hydrography of the Arkona Basin resembles the conditions in the Bornholm Basin more than those of the Danish Straights and the Belt Sea in SD 22 (Matthäus and Franck, 1992; Lass and Mohrholz, 2003), with pronounced thermohaline stratification and stagnation in the deepest areas of the basin. Spawning areas of western Baltic cod are in the deep, saline waters below 20 m, depending on area topography (Hüssy, 2011). The highly variable hydrodynamic conditions and the fact that cod eggs float in the water column cause their entrainment by currents, and their destination is determined by the prevailing winds and currents. Salinity limits the east–west exchange of eggs as a consequence of the stocks' differential requirement for neutral buoyancy. Superimposed on this, oxygen content and temperature have a significant effect on fertilization, egg/larva development, and survival (Hüssy, 2011). The long-term analysis of environmental conditions allowing survival of western Baltic cod eggs indicates that favourable conditions predominantly occur during the late spawning season in April/May. However, during the western Baltic cod stocks main spawning season in January to March, the suitability of the Arkona Basin for survival of this stock's eggs is limited, owing to the low temperatures often prevailing at that time of the year (Köster *et al.*, 2017). Unsuitable periods and habitats exhibiting the highest mortality rates are thus exclusively characterized by ambient water temperatures below the critical survival threshold. Despite the strong influence of water temperature on habitat suitability, the impact of habitat suitability on recruitment could not be clearly defined, suggesting that other mechanisms regulate year-class strength (Hüssy, *et al.*, 2012).

3.5 Stock Assessment data

3.5.1 Catch, Recreational fishery– quality, misreporting, discards

Two genetically distinct stocks are present within the management area and this split is accounted for in the catch matrix (3.6.2.). Recreational catches from Germany have been included in the assessment of this stock since 2013 and at this benchmark it was decided to also include Danish and Swedish recreational data. Discard estimates are available from Denmark, Germany, Poland, and Sweden have been included in the stock assessment for several years.

3.5.2 Relative proportions of eastern and western cod in SD 24

Stock splitting is based on otolith shape in combination with genetics or spawning cod sampled during the respective spawning time in SD 22 or SD 25. In recent years otolith shape analysis has developed into a useful tool for stock identification purposes (Campana and Cassleman, 1993; Bolles and Begg, 2000; Cardinale *et al.*, 2004; Mérigot *et al.*, 2007). Stock-specific otolith shape description based on Elliptic Fourier Analysis provides a means for classifying individuals caught in a mixed-stock area to their respective natal stocks. For Baltic cod, this approach has recently been documented as a potential tool to separate individuals belonging to the eastern and western stock (Paul *et al.*, 2013). This approach has been further developed and tested using genetically validated Baltic cod (Hemmer *et al.*, 2019) (Figure 3.5).

Stock splitting proportions are calculated separately for subareas 1 and 2 (Figure 3.1), due to an east-west gradient in stock mixing proportions (Hüssy *et al.*, 2016b). Three different approaches are currently used for stock splitting in SD24, all based on otolith shape analyses combined with genetics and spawning individuals, though with methodological differences.

The time-series of estimated proportions of eastern and western Baltic cod within SD 24 were available for the years shown in Table 3.1, based on otolith shape analyses from Germany and Denmark. Systematic differences in the proportion of mixing were found by subareas within SD 24, with a larger proportion of eastern cod closer to SD 25 (Figure 3.1 and Figure 3.5). The proportions of mixing in the easternmost rectangles in SD 24 and those in the middle of SD 24 were relatively similar (Figure 3.1). Therefore, these data were merged. The final proportions for splitting populations in SD 24 were estimated separately for two subareas, marked as Area 1 and Area 2 in Figure 3.2.

To prolong the commercial dataseries of stock mixing proportions back in time, historical survey data from 1985–1995 were used. From 1996 onwards, only commercial mixing proportions were used. However, there were several years without splitting data (1987–1991, 1997, 1999, 2001–

2004, 200–2007, 2009, and 2012). The missing information for single years, when the data for adjacent years were available, was filled by averaging the data from neighbouring years. To fill the gap in the Danish data from 2000 to 2008, the population splitting keys were derived assuming a linear increase in the proportion of eastern cod in the period from 1996 to 2013, both in Area 1 and in Area 2, the regression being based on the years for which data were available. The resulting proportions of western cod in SD 24, by years and subarea for 1985–2017 are shown in Figure 3.4 and Table 3.1. For a more detailed description, see WD2.

Three different stock splitting methods have been applied to the data used in the assessment as described below.

Method 1

This method has been used since the last benchmark in 2015. The methodology used to identify relative proportions of EB and WB cod in Danish commercial catches in 1996–2017 is described in Hüssy *et al.* (2016 a and b). The stock splitting proportions in Danish commercial data are available from 1996 onwards, however with several years of gaps in the time-series, 12 out of 22 years (1996–2017) (Figure 3.2). The baseline samples used in these analyses include both genetically validated fish, and fish for which stock origin was defined based on spawning activity at the time and in the area of either eastern or western cod.

Method 2

At the benchmark data compilation meeting in October 2018, the Thünen Institute (DE) presented an alternative stock splitting approach using a balanced and genetically validated otolith baseline with a good spatial coverage in SD 24 and adjacent areas, which allows the individual assignment of unknown cod otoliths to their stock of origin. This method was used to derive historical mixing proportions of WB and EB cod from samples, originating from German trawl surveys (1977–1995) and German commercial catches (2005, 2010, 2015–2016, active gear only; passive gear ratios are currently not considered). For stock proportions from surveys, only cod above 30 cm in length were considered. Details on the methodology are described in WD10.

Method 3

At the benchmark meeting in 2019, a third method for splitting Danish catches to stocks was introduced (WD11), and proposed to replace Method 1 in future (from 2018 onwards).The new method is a single coherent statistical model correcting for the effects of fish length, season, and yearly environmental changes while estimating mixing proportions. The method is general and can include any covariate suspected to effect otolith shape. It is not limited to the currently included covariates. The present new method 3 uses maximum likelihood to estimate otolith shape, otolith shape effects and stock mixing proportions in a single coherent analysis. Consequently, confidence intervals incorporating directly all data sources are provided. At the benchmark meeting, evidence of different effects on otolith contour shape was presented, including effects of fish length, season along with yearly variations in otolith shape. For example, year effects could potentially also be associated with sampling design. Further, it was shown that ignoring these – or other important - effects will lead to incorrect mixing proportion estimates. The new method (3) has been tested through simulation studies, and the adequacy of model fit to data is validated by residuals. The details of the method are described by Albertsen in WD11.

Decisions taken at the data compilation workshop

During the data compilation workshop it was decided to maintain the Danish historical stock splitting proportions derived with Method 1 for the years 1996–2017. These were decided to be supplemented by stock mixing proportions of German commercial data (only active gears in years 2005, 2010, 2015–2016) and German bottom trawl survey data from 1977–1995, based on Method 2. For years where data on stock mixing proportions were not available, interpolations were applied (averages of adjacent years) (Figure 3.4).

Mixing proportions were applied for the western Baltic cod covering the period from 1985 to 2017. The reason for not using the full period with available data (back to 1977) was due to inconsistencies in the SOP calculations and uncertainties in the official landings in the years 1981 to 1984.

Moreover, it was decided that Germany provides annual mixing proportions in SD 24 based on commercial samples, which are representative for German catches from active gear fisheries and based on samples from the German BITS in quarter 4, using the stock separation method described in WD10.

Additional decisions taken at benchmark meeting

At the benchmark meeting, the stock splitting Method 3 was evaluated. Method 3 is an improvement to Method 1 and was adopted by WKBALTCOD2 to be used to derive proportions of EB and WB cod in SD 24 from Danish commercial data from 2018 onwards. At the next benchmark, the stock splitting methods should be re-evaluated.

Suggestions for future work

The new stock splitting method (Method 3) revealed that strong year effects in otolith shape potentially exist. The reasons for the year effect could be limited spatio-temporal coverage of the commercial sampling or real changes in the otoliths between years. It was discussed during the meeting to increase the coverage and sampling level for commercial catches in SD 24 to be able to monitor the possible effects of the strong 2016-year class in order to use this year class as a "live experiment" to test if stock proportions changed. The hypothesis would be that with a very strong 2016-year class, the proportion of WB cod would increase in SD 24 compared to EB cod. If no such changes in proportion can be traced, it could be considered to use a 3–5 year mean in the mixing proportion instead of having an annual mixing proportion. Thus, a mean would reduce the noise and uncertainty, which may currently be introduced if the sample coverage is not very high or unrepresentative.

3.5.3 Catch data preparation

Landings in tons

Landings in tons by SD for 1985–2017 were obtained from WGBFAS reports. Total landings in SD 24 were adjusted to include only those representing the WB cod population. For each country, the relative proportion of cod landings in subareas 1 and 2 within SD24 were derived from national data. For earlier years, where this information was not available, extrapolations of the landings distribution from more recent years were applied. The weightings represented relative proportions of Danish, German, Swedish and Polish (main part of fisheries in SD 24) commercial cod landings taken in Areas 1 and 2. The landings in rectangles 39G2, 38G2 and 37G2 were used as representing Area 1 and landings in rectangles 39G3, 38G3, 37G3, 39G4, 38G4 and 37G4 were used as representing Area 2. The landings by rectangle from 2003 onwards were available from the STECF database (http://datacollection.jrc.ec.europa.eu/dd/effort/graphs). Danish landings by rectangle back to 1994 were derived from the national database. The relative distribution of German landings between Areas 1 and 2 from 1994–2002 was set to the average of that in the years 2003-2013. The total landings of Germany, Denmark, Poland, and Sweden in SD 24 (derived from earlier ICES WGBFAS reports) were used as weighting factors to derive an average distribution of landings between Areas 1 and 2 separately by country for Denmark and Germany and for the remaining countries, the information was combined. These average proportions of landings between Areas 1 and 2 were then used as weighting factors to derive an average splitting key for landings in SD 24 (from the two separate stock splitting keys for Areas 1 and 2). The resulting landings of WB cod population by SD is shown in Figure 3.4.

Landings at age

Prior to the benchmark meeting, a small calibration exercise (49 sliced otoliths from BITS Q4 2017 and 2018 and commercial samples from 2018, covering age 0-10) was coordinated by Germany, which identified still differences in age readings in SD 22. The overall percentage agreement was 85% and overall average percentage error of 9%. Unlike previous cod otolith exchanges, this was the first exchange, which could consider objective evidence from age-validated otoliths (from a mark-recapture study using tetracycline marked wild cod in SD 22) and high-resolution length–frequency distributions of juvenile cod caught in poundnets (see McQueen *et al.*, 2018 and WD18). Although the overall level of agreement was high, there were issues of counting the first true winter ring and misinterpretation of the edge due to the month of recapture by two countries, resulting in overestimation of age by usually 1 year in 21 out of 49 otoliths. However, two of the countries do not read sliced otoliths, as was used in the exercise, on a regular basis. Cut otoliths are read with reflected light while sliced otoliths are read using transmitted light, hence the appearance of the visible ring patterns is reversed between the methods. This methodological difference between countries persists since 2008 and was already an issue during the last cod otolith exchange prior to the 2015 benchmark (where sliced otoliths were also used).

It was decided to have a skype meeting (DK, DE and SE) prior to WGBFAS 2019 to discuss the age-reading issues and if needed to correct age estimates before the assessment working group. The Thünen Institute offered to slice and take images of all (or of a significant subset of) commercial and survey otoliths of DK and SE from SD 22-23 from 2019 to compile a dataset for a thorough between-country age reading comparison. If the effort is not too high, this could be repeated each year to ensure high quality age data of this shared stock. It was decided during the benchmark to consider age information from otolith readings for WB cod (SD 22-23) to be of sufficient quality to be used in stock assessment.

In the previous assessment the age structure from SD 24 was not used because i) a large part of the stock in SD 24 has been identified as eastern Baltic cod with possibly different age structure than the western population, ii) given the large proportion of eastern cod found in SD 24, the age information for this area is considered uncertain due to age reading problems of EB cod otoliths in later years.

In previous assessments (WGBFAS 2018), data on age or size structure of cod catches in SD24 have not been used. Instead, the catch-at-age in SD22 has been raised to account for tons of WB catches in SD 24. Similarly, catch-at-length in SD25 has been raised to account for EB cod catch in SD24. This assumes that the WB cod in SD 24 have the same age structure as the WB cod in SD 22, and the EB cod in SD 24 have the same size structure as the EB cod in SD 25. At the benchmark data meeting 2018, this assumption was evaluated using data on observed length distributions in SD 24 for the period 2000-2013 (compiled in Intercatch). This was done by comparing the calculated catch at length in SD24 with the observed catch-at-length. The calculated catch at length was obtained by summing up i) the WB cod fraction of catch at length in SD 24, obtained when raising catch a length in SD22 by the WB cod tons from SD 24; and ii) the EB cod fraction of catch at length in SD24, obtained by raising catch a length in SD25 by the EB cod tons from SD 24. The exercise showed that the calculated and observed catch at length in SD24 were relatively similar (Figure 3.12), indicating that the assumption used in the assessment is reasonable. Therefore, the approach previously applied in the assessment was suggested to be continued.

Decision taken at the data compilation work shop

Landings at age for the entire western cod population (i.e. including landings in SD 24) were obtained by upscaling the landings at age in SD 22 by the ratio of landings of WB cod taken in SD 24 compared to SD 22. Landings at age in SD 23 were subsequently added, to get the landings-at-age of WB population for SD 22-24. Thus, the age structure of the landings in SD 24 was assumed similar to that in SD 22. The age structure from SD 23 was not included in distributing landings from SD 24 to ages, due to different fishing patterns in SD 23 (trawling ban).

Discard catch-at-age

Discards for the period 1996-present was calculated the same way as in previous assessments: Discards before 1996 where no discard data have been available were extrapolated by an average discard ratio from the period 1999–2003. For more details see WD13.

Recreational catch

Until this benchmark only German recreational catches have been included in the assessment. At this benchmark also Danish and Swedish recreational data were included and the German recreational data were updated back in time (Figure 3.13).

All recreational cod catches taken in SD 22-24 were considered to be WB cod. The recreational catch-at-age data from the three countries were merged and the weight-at-age data were combined and weighted with the catch by country.

Different approaches to estimate catch-at-age were used by SD. In SD 23 only Danish and Swedish recreational fishery were considered for stock assessment as the two countries have the majority of the recreational catch in this area. In SD 22 and 24 only German and Danish recreational catches were considered.

Reconstruction of recreational catches were based on agreed expert assumptions.

Annual Catch

Swedish annual catches from SD 23 for the tour boat fleet were derived from an onsite survey program and logbooks for the period 2011–2017. The 2017 value also estimated catches from private fishing boats, which was added to the tour boat catches. The estimated amount of Swedish recreational catches varied between 80–200 t during this period. Catch data between 1994 and 2011 was estimated from a model taking into account nine different interviews with skippers on their historic catches combined with the wind and temperature information from two harbours located within the area. Recreational catches from 1985–1994 were estimated as an average of the data in 1994–2011. During the WKBALTCOD2 meeting it was decided that although this historic approach from 1994–2011 gave a relatively stable catch level of close to 150 t a year, the model is based on relatively few years with real data. It was acknowledged that this approach could be used in future calculations. However, it was considered more realistic to use an average for the catch data in the period with sampled data (2011–2017) for the historic time-series (1994–2011).

Danish annual catch data from SD 22, SD 23 and SD 24 derived from 2- annual Statistics Denmark recall survey (2009–2018) scaled to the observed value from the on-site studies in SD 23 in 2016–2018. For the period 1985–2008, Catch per year has been calculated as the mean catch per year for the period 2009–2018, weighted for each year with the number of Danish citizens being 18–65 years old (age range for which holding a fishing license is mandatory).

German annual catches from SD 22+24: Face value annual catch data were used from 2005–2017. Previous years (1980–2004) were based on the average catches from 2005–2017 (Table 3.2).

Length

SD 23: Length composition from annual onsite sampling was used in combination with Danish and Swedish data for the years 2012, 2013, 2016, 2017, and 2018. An average of the time-series was used for historic data (1985–2012).

SD 22+24: Face value annual length data were used from 2005–2017 from the German sampling program. Previous years (1980–2004) were based on the pooled length distribution from 2005–2017.

Age

In SD 23: Face value age data used in 2016 and 2017. ALK and WECA were calculated from BITS/commercial data from the Swedish and Danish sampling programs.

The Swedish age were derived by converting sizes to ages using an age–length key for each individual year. R package {Fishmethods} was used to calculate ALK type 2, i.e. proportions-at-age per size.

In SD 22+24: Estimation on German recreational cod removals (harvest and dead releases) in numbers were distributed according to the recreational length distribution and then matched with the ALK from commercial or BITS data for each year. Age proportion per length was used to group numbers-at-length into numbers-at-age categories. WECA was calculated using the length mass coefficient from the same data sources.

Release (discards)

Recreational catch data are collected for both harvested and released cod. A large amount of cod is released alive voluntary due to minimum landings sizes and bag limits (Ferter *et al.,* 2013). Anglers release fish with the assumption that they will survive. Several studies exist that have estimated post-release mortality of released cod (Capizzano *et al.,* 2016; Ferter *et al.,* 2015; Weltersbach and Strehlow, 2013). The amounts of dead recreational releases was accounted for and included in the recreational catch data. In Germany 100% of land-based releases were considered dead (precautionary approach as no studies are available). Dead sea-based releases were estimated applying a 11.2% post-release mortality rate to all sea-based releases based on a post-release mortality study from 2012 (Weltersbach and Strehlow, 2013). German annual release data were used based on face values from 2009–2017. Previous years (1980–2008) were based on the average from 2009–2015. In Denmark, all releases were considered sea-based and the same post-release mortality rate applied.

For more information about the national recreational catches collection and estimations see, WD14, WD15 and WD16.

3.5.4 Weights, maturities, growth

Mean weight at age in the catch

Annual weight at age in landings in SD 22 was available from 1994 onwards, from earlier WGBFAS reports. For earlier years, average values from the earliest period 1994–1998 were applied.

For SD 23, annual mean weight was available from 1997 onwards. For earlier years, average values of the period 1997–1999 were applied.

For SD 22-23, data were available directly from InterCatch for the years 2014 onwards. For earlier years, the mean of SD22 and SD23 was applied, weighted by landings-at-age. These data were used to represent mean weight in landings of the entire WBC stock.

Weight-at-age in discards in SD22-23 were available for 2014 onwards. For earlier years, values equal to 2014 were applied. These were used to represent mean weights of total discards of WBC stock.

Total weight-at-age in catch is derived by averaging the mean weights-at-age in landings, discards and recreational catch, weighted by respective catch numbers.

Mean weight at age in the stock

Weight-at-age in the stock for ages 1-3 were calculated from BITS Q1 data for SD 22-23, following the same calculation procedures as previously used for cod in SD 22-24 (see stock annex from 2015). For ages 4+, mean weight in the stock was set equal to the mean weight in the catch.

Maturity ogive

Maturity and spawning probability were estimated from BITS Q1 data for SD 22-23, as SD24 consists of a mix of eastern and western Baltic cod.

Spawning probability separates between the gonad development stages where the fish is likely to spawn or skip spawn, while proportion mature includes all the fish that are mature. Especially for younger ages, there are large differences between spawning probability and proportion mature, thus spawning probability was chosen to be used for ages 1-4. Due to very few older fish in the samples, maturity/spawning probability was set to 1 for ages 5+.

In the previous stock assessment, proportion of spawners has been used as a 3-years running mean, e.g. the values for 2017 used in assessment were based on the average of the estimates for 2015-2017. A 3-years means has been applied due to a relatively large variability of the data. As the 3-year running mean still contains considerable interannual variability of the data to be used in stock assessment, the WKBALTCOD2 data meeting agreed to use a 5-year running mean instead, that captures the long-term trends but further reduces interannual fluctuations (Figure 3.14)

SDs are defined based on coordinates given for individual fish data in DATRAS:

SD 22: ShootLat>53.5000 & ShootLat<=56.0000 & ShootLong>9.5000& ShootLong]<=12.0000
SD 23: ShootLat>=55.5000& ShootLat<=56.0000& ShootLong>12.0000& ShootLong<13.0000

The calculation of proportion of spawners in the population was scrutinized to ensure consistent treatment of data in the entire time-series. This resulted in some revisions to the time-series used by WGBFAS previously, though the overall calculation procedure is unchanged and the revisions do not affect the overall dynamics in maturation. Part of the revision is due to the area now being defined by the above given coordinates, while it previously was based on statistical rectangles. Figure 1 in WD17 shows the comparison on annual proportions of spawners in the new revised time-series compared to the one used previous in WGBFAS.

3.5.5 Survey data

Different options for calculating survey indices for WB cod, taking into account stock mixing, were explored. In previous assessments since the last benchmark, survey data east from 13 degrees longitude in SD 24 were not used given the lack of mixing proportions. For WKBALTCOD2 the German stock separation method (Method 2) was used to determine the individual stock affiliation of samples from the German BITS, covering all years in the period from 1992 to 2017 for quarter 4 and selected years for quarter 1 (1995, 2000, 2005, 2010, 2015–2016). Only individuals with a minimum size of 20 cm were analysed. This time-series of individual stock assignment was included in survey modelling by C. Berg (Berg and Kristensen, 2012) to provide east-west population-specific estimates of relative abundance of WB and EB cod in SD 24 (WD2).

CPUE by age and by length for SD 22, 23, and 24 are available from ICES DATRAS database (BITS-Q1 and BITS-Q4). To account for mixing with the eastern Baltic cod in SD 24, different options were considered for calculating survey indices, i) 13 degree, ii) Hard borders in SD 24, and iii) Soft border from SD22-26. The three options were discussed during the data compilation workshop and reflect different approaches to account for stock mixing. The 13 dregree approach was also used in the SPALY assessment, and therefore the only adjustment at WKBALTCOD2 was the prolonging of the time-series that earlier started in 2001 to start in 1996 (Q1) and 1999 (Q4). i) The 13-degree approach assumes that all cod west of 13 degrees longitude are of western origin and those east of 13 degree belongs to the eastern Baltic cod stock. In the second option ii) Hard borders in SD 24, the distribution of east/west cod within SD 24 is modelled but not allowing western cod to be east of SD 24 (i.e. in SD25-26) and eastern cod in SD 22. In the last model iii) Soft border from SD22-26, both eastern and western cod can mix in the entire area of SD22-26. However, currently there is little genetic evidence of stock mixing beyond SD24, and the mixing beyond SD24 in this exercise in purely based on model extrapolations.

The results from all three models were compared for internal consistency to investigate whether any of the approaches modelling the stock split in SD24 improved the consistency of the survey index compared to the SPALY option (13 degree split).

i. 13 degree. (Figure 3.15a, 3.15b and 3.15c);
ii. Hard borders around SD 24 (Figure 3.16a, 3.16b and 3.16c);
iii. Soft border around SD 24 (Figure 3.17a, 3.17b and 3.17c);

The internal consistency of survey indices did not improve for survey models ii) "Hard" or iii) Soft for quarter 1 and, in fact, it became much worse for quarter 4 compared to the original survey model applying 13 degree split.

The decision at the WKBALCOD2 was to keep the original survey model, i.e. allocating all cod west of 13 degrees longitude to the western stock and all cod east of 13 degrees to the eastern stock. This is due to the best internal consistency of this survey index among the ones investigated. Furthermore, the 13 degrees split approach to the survey is most consistent with the commercial data in terms of dealing with stock mixing. In the commercial data, it is assumed that the western stock in SD24 has the same length/age structure as in SD22, which is similar to the 13-degree split approach in the survey. In contrast, the other approaches to account for stock mixing in survey indices allocate fish to stocks by length, resulting in different assumptions concerning length/age structure of stocks in SD24 than those used for commercial catches. Future work in this area should consider a common framework, where stock mixing can be accounted for both in commercial catch and in the survey in the same way, allowing to explore different mixing assumptions, while maintaining consistency between catch and survey data.

Poundnet survey

A new survey, targeting juvenile cod (<38 cm, age groups 0 and 1) was presented at the data compilation workshop. The survey is conducted in cooperation with German poundnet fishers operating in shallow coastal waters around Fehmarn Island, and provides an 0-group abundance index from 2011-present (Figure 3.18). It was decided during the data compilation workshop to test if the age-0 abundance index from the survey could be included in the stock assessment model, although it currently consists of a relatively short time-series and only covers a small area of the stock distribution area. The biological reasoning for including the poundnet survey was that while the trawl survey covers areas in the western Baltic Sea deeper than 10 m (ICES, 2017), juvenile cod are reported to inhabit shallow inshore waters (Pihl and Ulmestrand, 1993) and may preferentially occupy shallow-water vegetated habitats (Freitas *et al.*, 2016) which are not adequately covered by the BITS.

Additionally, the scientific surveys are often criticized by fishers, as these surveys only cover a very short period. This poundnet survey has a four-month duration (samples are collected throughout September to December) and is therefore considered to provide a robust estimate. Sampling of the poundnets is planned to continue, so the time-series will continue to be extended. In future, the age-1 abundances estimated from the poundnet sampling may also be considered for the stock assessment. For further details, see WD18.

3.5.6 Assessment model

The state space stock assessment model (SAM) was used as in previous years. No other stock assessment model was applied.

3.5.7 Exploratory assessment analyses

Model Settings

A residual pattern was evident for the oldest age class in the survey and for the last years in the catch matrix (Figure 3.19). The residuals in the surveys for older ages were improved by decoupling of the all age groups in the survey catchability parameters. In the SPALY settings, the two older age groups were coupled. This also improved the log likelihood from-429.9 to -412.0 and the AIC from 899.9 to 868.1.

Further, correlations across ages, as an effect towards a year effect in the surveys was tested for improvement in the model. For the quarter 4 survey this had a small positive effect on the log likelihood however, the retrospective pattern for recruitment was worsened and therefore this setting was not applied. For the quarter 1 survey the correlation across ages improved the log likelihood from -412.0 to -393.8 and the AIC from 868.1 to 835.6. The retrospective pattern were similar to the former settings. Correlations across ages were therefore accepted for the Q1 survey by the benchmark group.

It was not possible to improve the residuals in the catch matrix during the conducted test runs. Several solutions were tested including catch multiplier and removing of survey data. The residuals are an effect of a disagreement between the catch matrix and the survey results (from 2014-2017).

The period coincides with decreasing quota, however the quality of the mixing data could also cause this inconsistency. To test this hypothesis an exploratory run was conducted with a catch multiplier for the years 2014-2017. The results were a catch multiplier between 60–100% in the year spanning 2014–2017 which were considered an unrealistically high amount of misreporting of landings(Figure 3.20).

The 4th quarter survey is considered to cover less reliably the stock distribution as the stock in some years at the survey time has not migrated into the depth were it is catchable for the survey. Therefore, a second set of exploratory runs was conducted to test if the residuals could be improved, by removing the 4th quarter survey was from the assessment. There was a slight improvement in the residuals for the older age classes however, the residual pattern was still there and it did not seem to be justified to remove a whole survey time-series for a slightly improved residual pattern for the older age groups.

The group therefore decided to keep both data sources in the model as long as there were no really good evidence of mistrusting either the survey or the catch data.

As a more complete recreational dataset was introduced and included in the assessment it was discussed if it was the most appropriate way to have the recreational and commercial fleet as a combined CANUM and WECA or if the model should be set up to run with 2 separate fleets.

The model was set up and an exploratory run was conducted with two separate fleets. With respect to SSB the two models (single fleet compared to two fleet model) gave a very similar results, although the SSB in the two-fleet model gave lower SSB , however within the confidence internals), in the period 1994–2002 as well as in the last year. F estimates were also relatively similar in the two models (Figure 3.21). F pattern by age showed a very different F pattern by age between recreational and commercial catches. F has been decreasing for in the commercial catches in the latest 8-10 years for all ages (except age group 1) whereas F has increased from 2010 to 2015 with a large decrease in the latest two years (Figure 3.22). The 2-fleet model did not improve the residuals, it got slightly worse and the retrospective plot got worse by the 2-fleet model with several years having values outside the confident intervals (Figure 3.23). Furthermore, a 2-fleet model is more complicated both as an assessment model but also in the advice and for short-term forecast and therefore the decision by WKBALTCOD2 was that, as the new model did not improve the fit, the final assessment was the 1-fleet model with the combined recreational and commercial catch.

3.5.8 Final assessment

The setting and input data used in the final assessment can be found at www.stockassessment.org with the assessment name WBcod_Benchmark2019. The model seems to fit relatively good to the catch data, however especially for age group 7 some years the observed catches have been somewhat lower than the model estimate (Figure 3.24). The model fit to the surveys can be found in Figures 3.25, 3.26 and 3.27. Residuals in the final settings show a similar pattern as was evident in the latest assessment with negative residuals in later year in the commercial catch matrix (Figure 3.28), however the pattern in the oldest age group in both Q1 and Q4 surveys has been improved by in the new assessment (Figure 3.28 and Figure 3.29). The selection pattern from the combined fleet showed that the fish is 90% selected at age 3 and fully selected at age 4 with a sigmoid selection curve (Figure 3.30). In the final model with the combined fleet F showed a decreasing pattern in F since 2008 for all age group except age 1, although the decrease has been largest for the older age groups (Figure 3.31). The leave one out plots indicates that without the Q4 survey SSB would be estimated lower and F higher than with the combined surveys, as the Q1 surveys has information from one more year than the other data point, and when this data point is left out information from latest year is missing (Figure 3.32). The retro plot for SSB gives a Mohn's rho at 0.12, -0.02 on F, 0.27 on R and 0.09 on the catch matrix. Although the levels are within acceptable limits, the retro in F shows a very large variation (although not in the same direction), and although the low Mohn’s rho there seems to be problems in a correct estimation of F (Figure 3.33). The final estimate for SSB and F is shown in Figure 3.34 and with the new prolonged time-series. It is evident that the stock had a historic low in 1991–1992 with SSB just below 10 000 t, and in relatively few years the stock increased to a historic high in 1997 with more than 40 000 t SSB. Currently the stock has been at a low level 2009-2016 but has in later years increased again close to the long-term mean (Figure 3.34). F has historically been very high for this stock with a mean F (3-5) at 1.2 in the period 1985–2002, then F decreased to around 1 in the next 10 years’ period 2003–2013, with a decreasing trend in later years.

3.6 Short-term projections

The following procedure was decided during the WKBALTCOD2 for the short-term projections. The short-term projections are simulated forward from the last year in the assessment. The last assessment year is the year the assessment is conducted in, because of the 1st quarter survey. In the last year of the assessment the estimates and their estimated uncertainties are used (including for recruitment), but for the following forecast years recruitment is sampled from the most recent 10 recruitment estimates (1000 times with replacement) (Figure 3.35). This constitutes a small change from the previous benchmark, where the short-term forecast was set to start one year prior to the last assessment year, but then still use the last assessment year's recruitment estimate.

The resulting difference between the two options is relatively small, and the implication for the short-term forecast is small, as most age groups relevant to the fisheries (+2) are informed with data in the short term. However, for communication, it is clearer to use the values from the summary table output as input in the short-term predictions.
Selection pattern and stock weight is used in the short term and it was decided to use the latest 3-year average.

3.7 Appropriate Reference Points

Biomass reference points

The stock recruitment relationship used included data from 1985-2017. WKBALTCOD2 considered six different stock characteristic types documented by ICES in "ICES fisheries management reference point for category 1 and 2". The stock recruitment plot did not indicate a clear S-R relationship, there is however some evidence of recruitment being impaired at very low spawning stock levels, though it was not possible to estimate a breakpoint (Figure 3.36).

As no breakpoint in S-R could be defined, WKBALTCOD2 decided by to use an average of the lowest SSBs in 4 years corresponding to good recruitment (the 1991, 1993, 2003 and 2016 year class). The average SSB corresponding to these 4-year classes was 14 535 t. It was decided by the group to round this number and hence, B_{lim} for the western Baltic cod equals 14 500 t. Using the ICES standard procedure this corresponds to a B_{pa} at 21 876 t (B_{pa} = 14500* EXP (1.645*0.25).

Fishing mortality reference points

F_{MSY} was calculated using ICES standard software EqSim. Stock–recruitment was defined using a hockey-stick function, setting the breakpoint at B_{lim} (14 000 t) (Figure 3.38). The entire time-series was used for S-R. For the biology and selectivity, average values from 2015–2017 were used, to account for trends in the time-series. F_{MSY} is relatively well defined for this stock (Figure 3.37), and was estimated at 0.26 (ranges F_{MSY} low= 0.18, F_{MSY} high= 0.43).
Precautionary fishing morality reference points were estimated to be at F_{lim}=1.45 and F_{pa}=0.99.

3.8 Future Research and data requirements Subarea

Mixing (model/ otolith/tagging/)

Regular updates on mixing proportion of eastern and western Baltic cod in SD 24 are required. Further, continued improvements to the otolith shape analyses used for stock separation and to the genetic baseline are expected. The possible differences in proportions of eastern and western cod by size and season should be further explored. Good geographical coverage of samples is required to derive improved, more representative stock splitting proportions.

Otoliths from commercial fisheries samples of Sweden and Poland from SD 24 should be included in future shape analyses.

3.9 Figures and tables

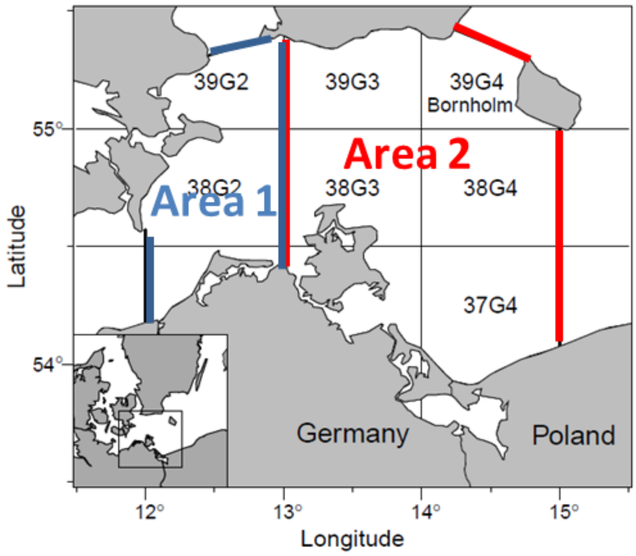


Figure 3.1. Map of SD 24 (mixing area of western and eastern cod) and subareas (Area1 and Area 2) for which separate mixing proportions were estimated.

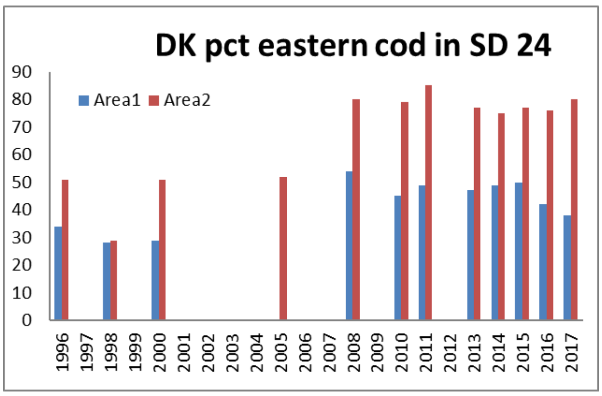


Figure 3.2. Proportions of EB cod in subareas 1 and 2, in years for which data are available from Danish commercial samples.

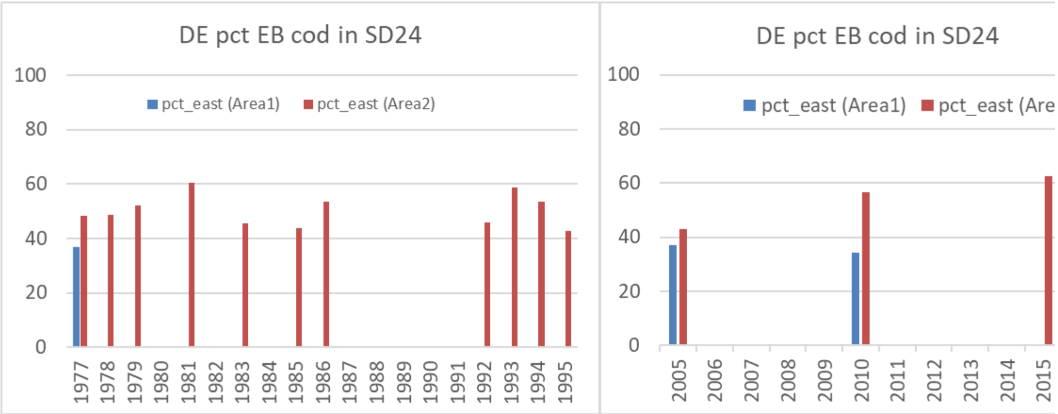


Figure 3.3. Proportions of EB cod in subareas 1 and 2, from German survey (left panel) and from German commercial samples from active gears (right panel).

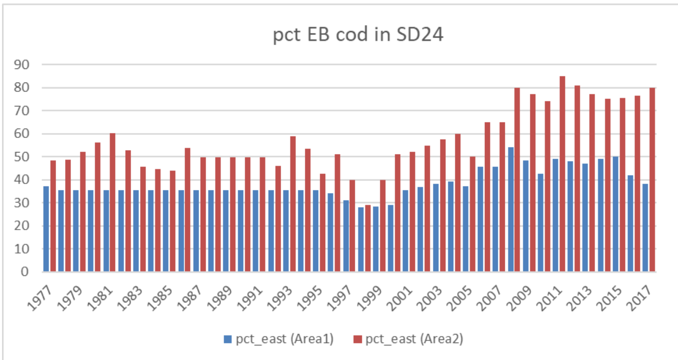


Figure 3.4. Proportion of EB cod in SD 24, by subareas, including extrapolation for years with missing data.

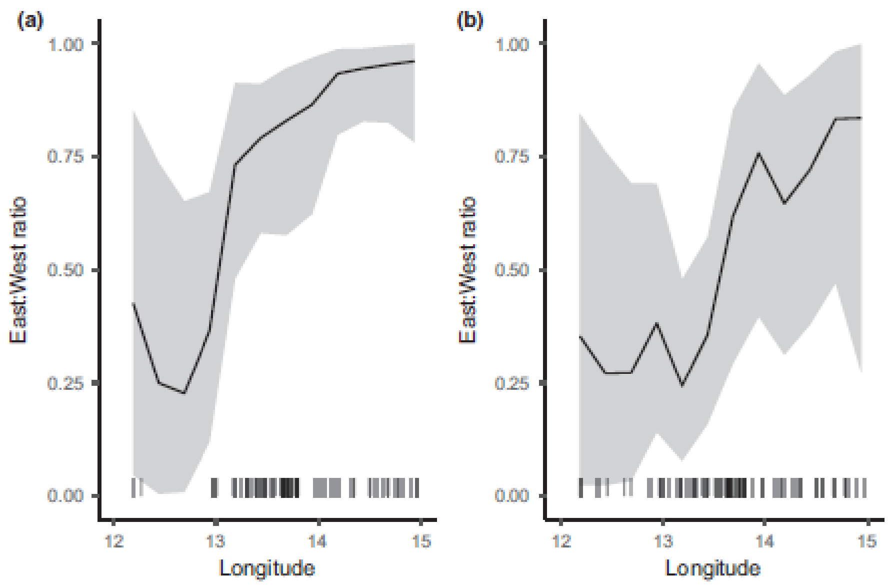


Figure3.5. Modelled geographical distribution of mixing proportions for adult cod (a) and by juvenile cod (b) by longitude in the Baltic sea (Hemmer-Hansen *et al.,* 2019).

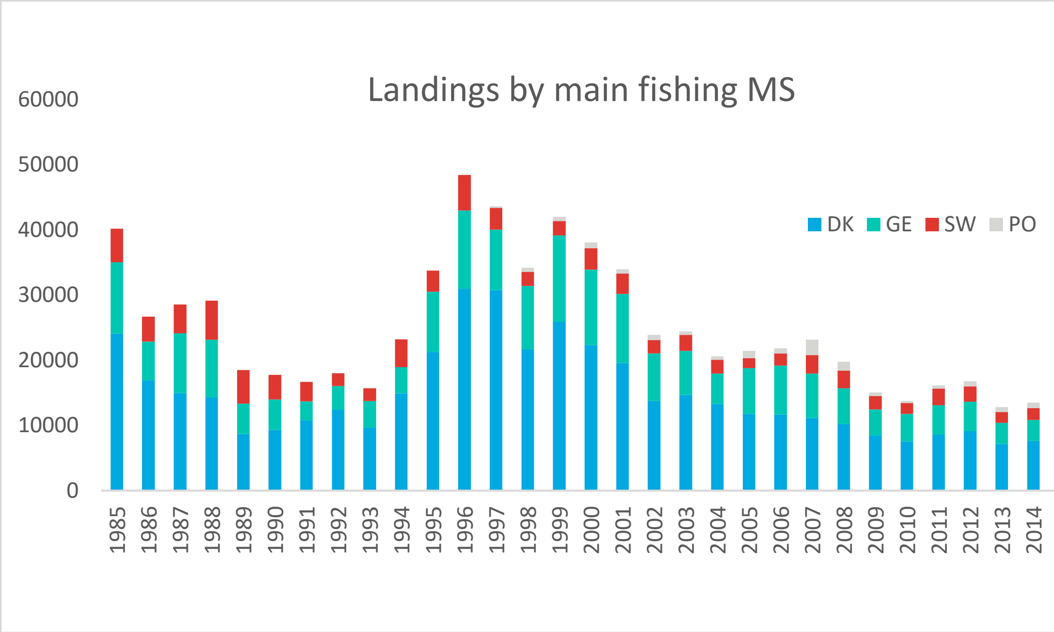


Figure3.6. Landing by Member state (MS) in the period from 1985 to present time from the 4 most dominant fishing nations on this stock.

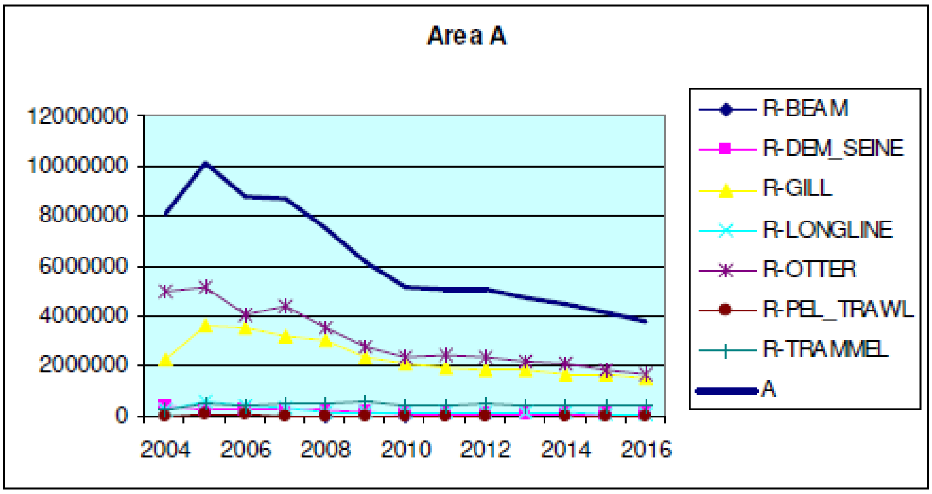


Figure3.7 Figure on effort in the western Baltic management area, from STECF report (2017).

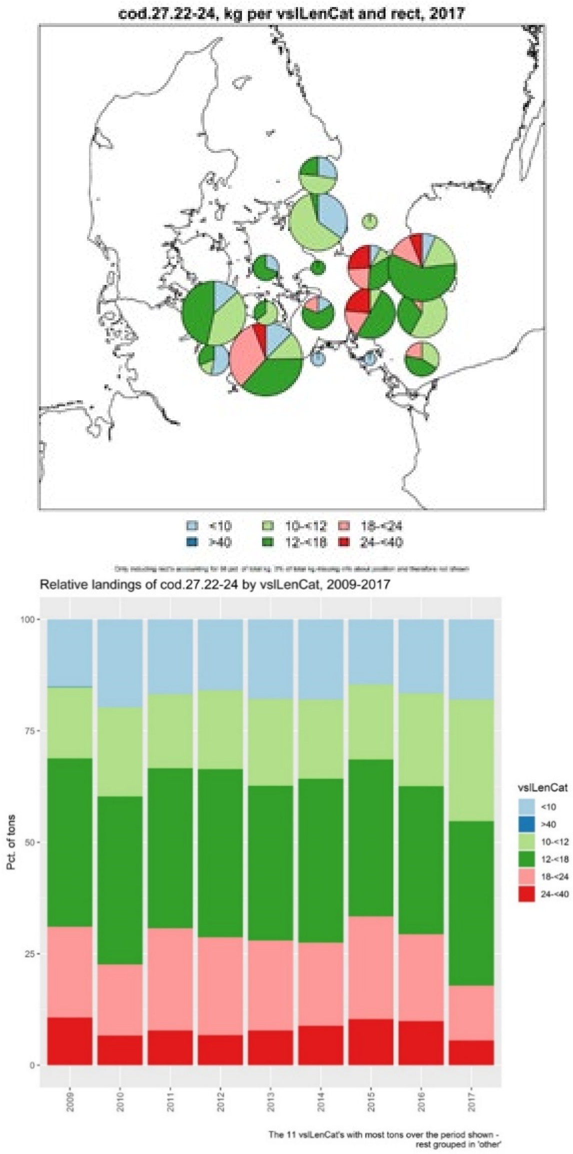


Figure 3.8. Left: Cod landings by ICES square and vessel length. Right: relative distribution of landings by vessel length and year.

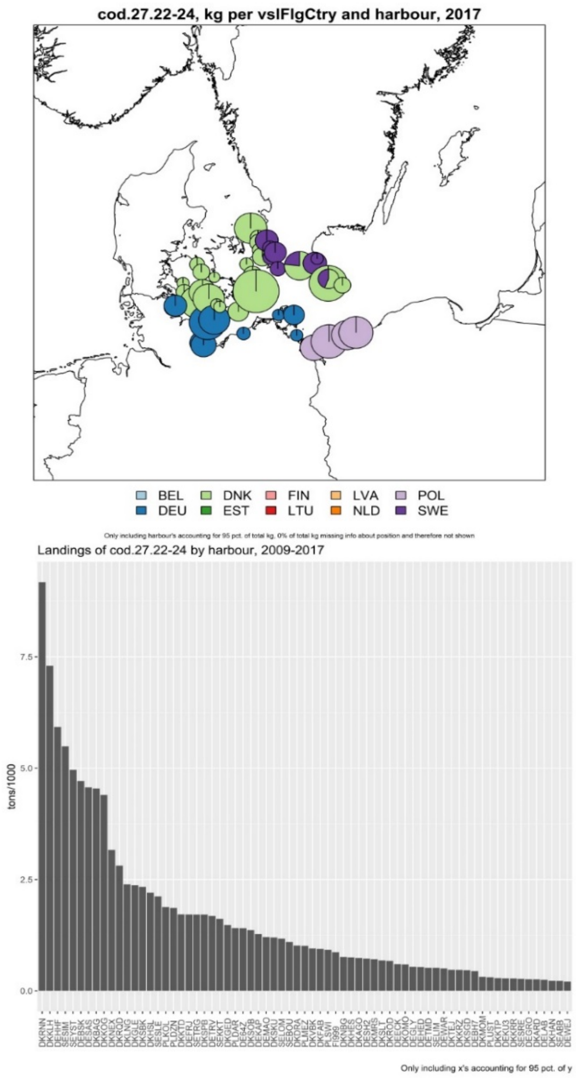


Figure 3.9. Left figure Landing by country and harbours. Right figure. Landings of western Baltic cod in amount by harbours.

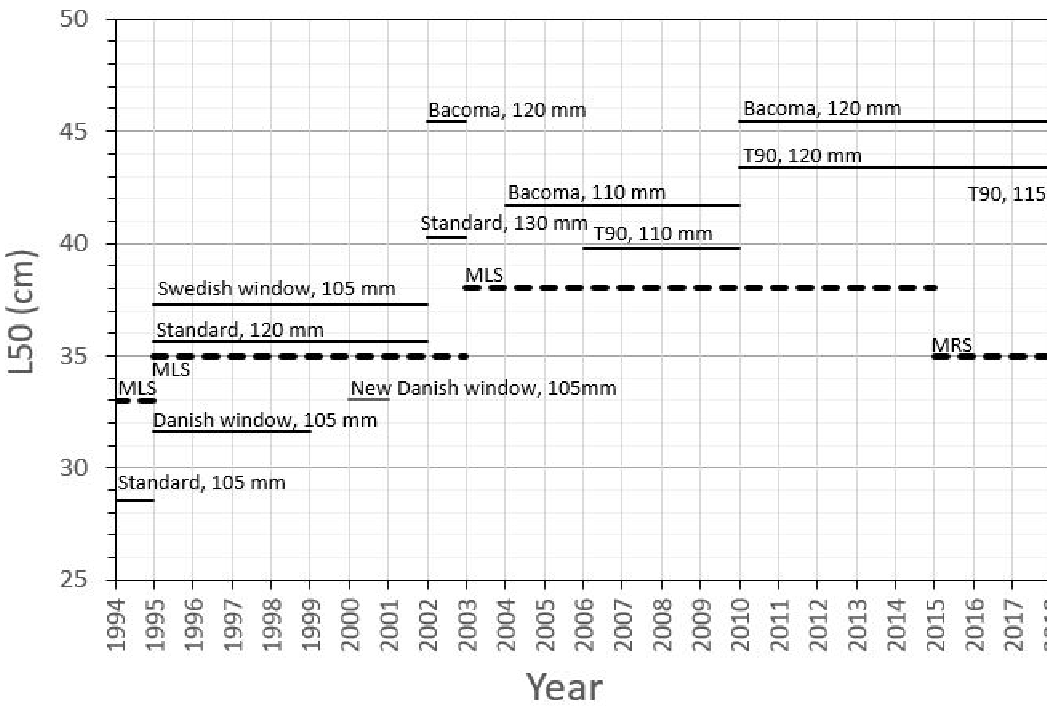


Figure3.10. Different management regulation in the time from 1994 to 2019. Figure is from Valentinsson *et al.,* 2019

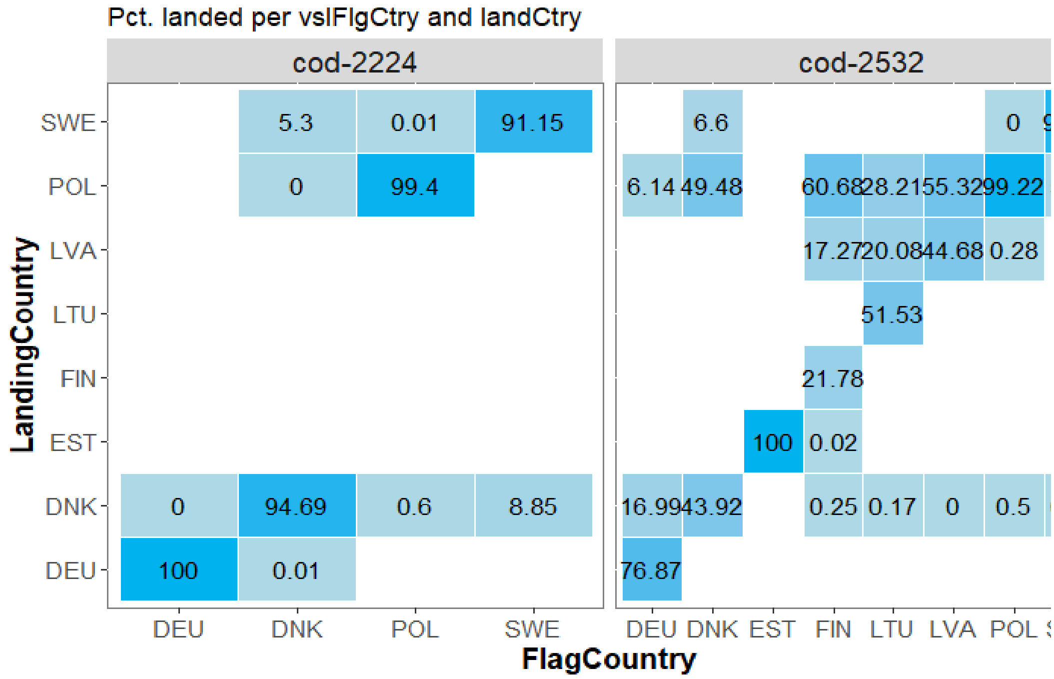


Figure3.11. Percentage of foreign landings by country in the western and eastern Baltic management area.

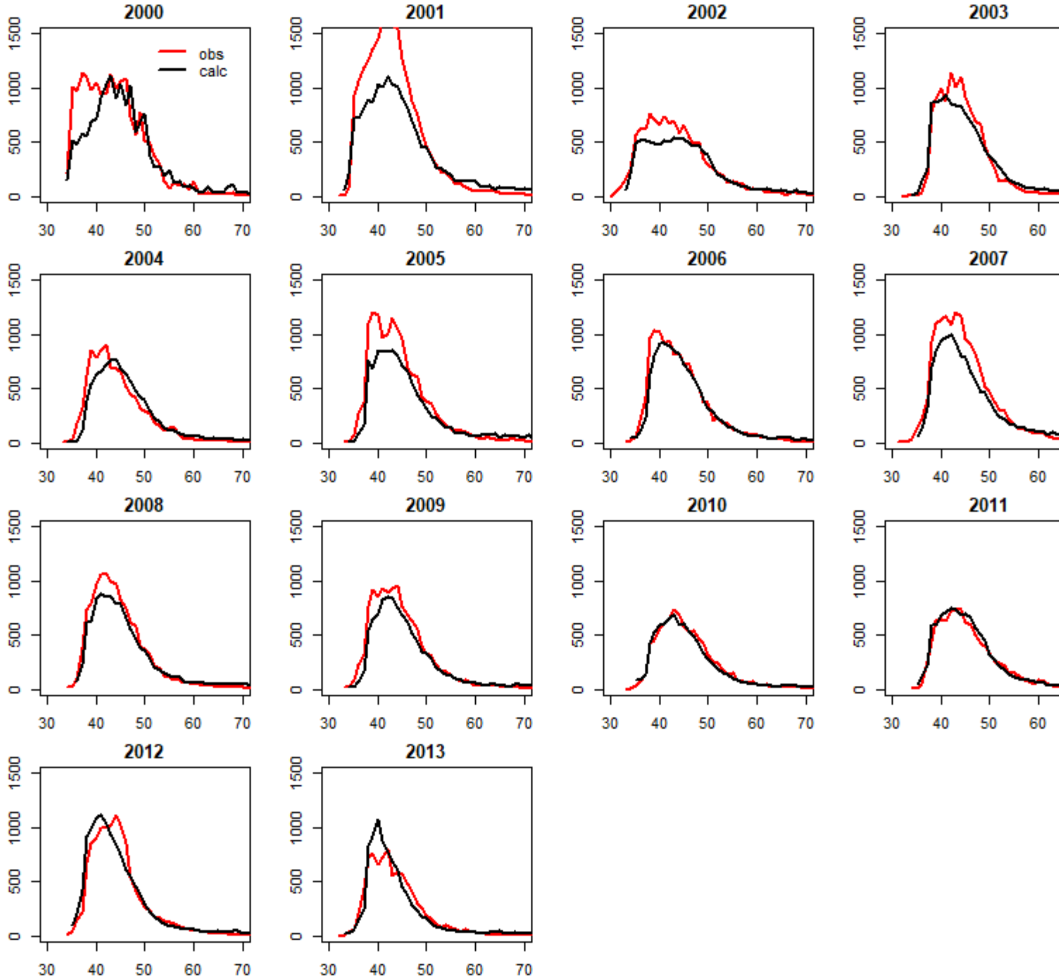


Figure 3.12. Observed and calculated landings at length in SD 24.

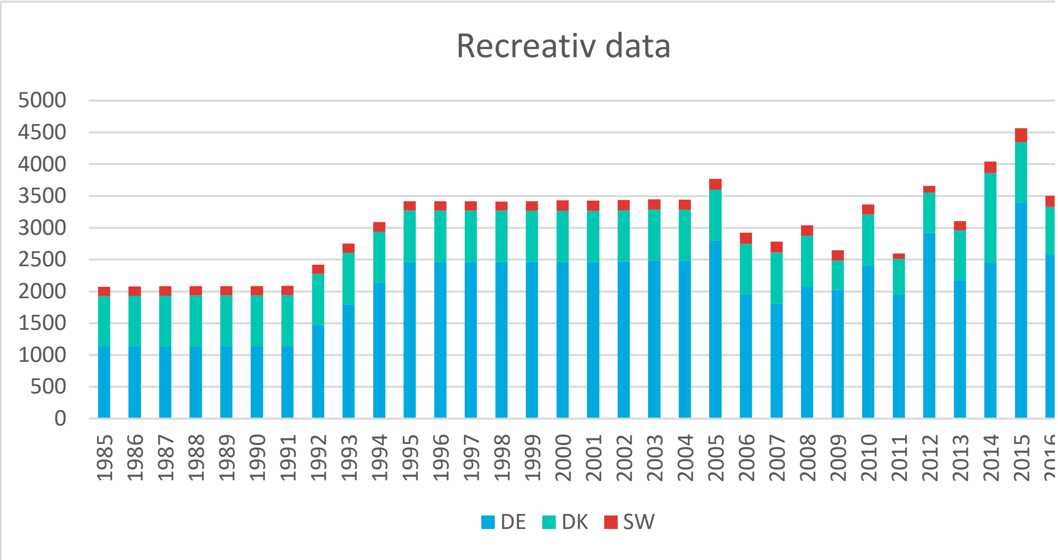


Figure 3.13. Recreative catch data of WB cod population by country.

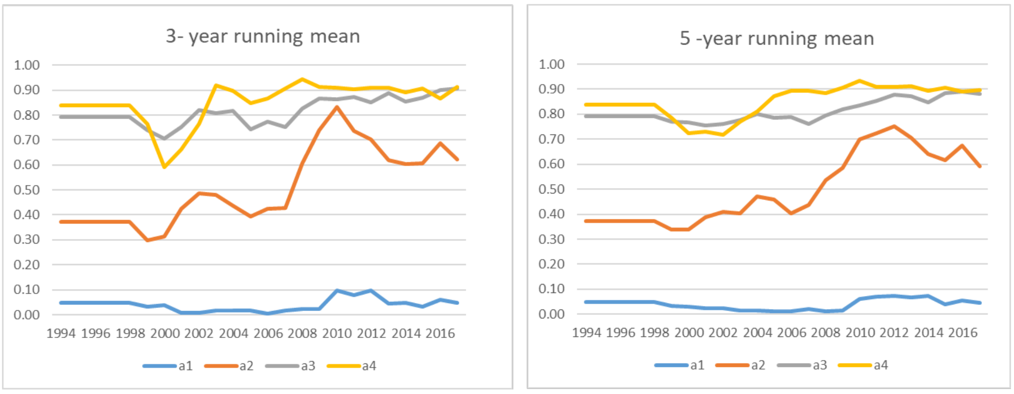


Figure 3.14. 3-year running mean proportion mature at age compared to a 5-year running mean.

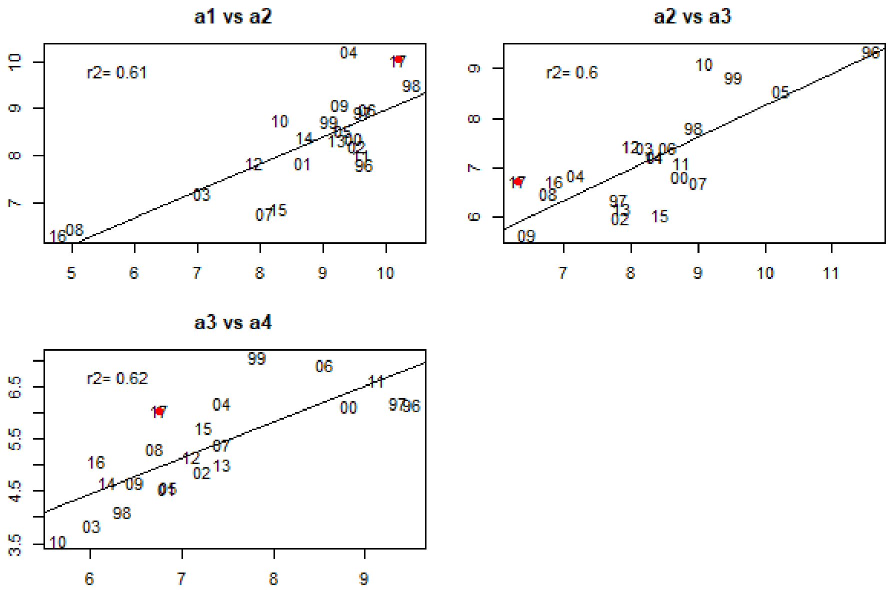


Figure 3.15a. Internal consistency of CPUE for survey model “13 degree” (BITS-Q1 data).

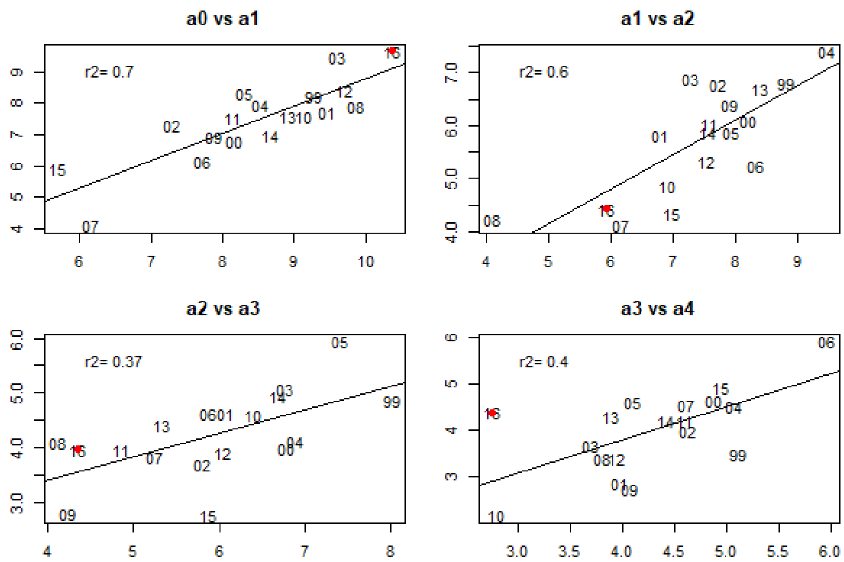


Figure 3.15b. Internal consistency of CPUE for survey model “13 degree”(BITS-Q4 data).

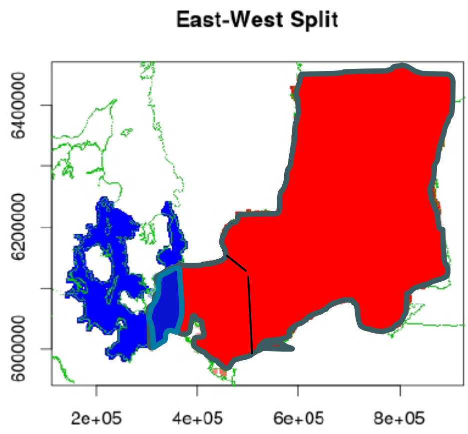


Figure 3.15c. Coverage area in the survey model "13 degree".

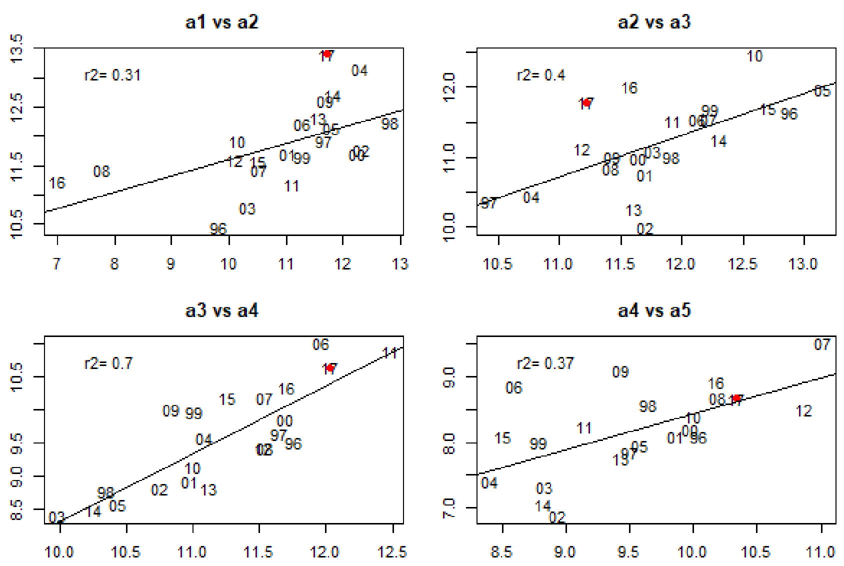


Figure 3.16a. Internal consistency of CPUE for survey model "Hard 24" (BITS-Q1 data).

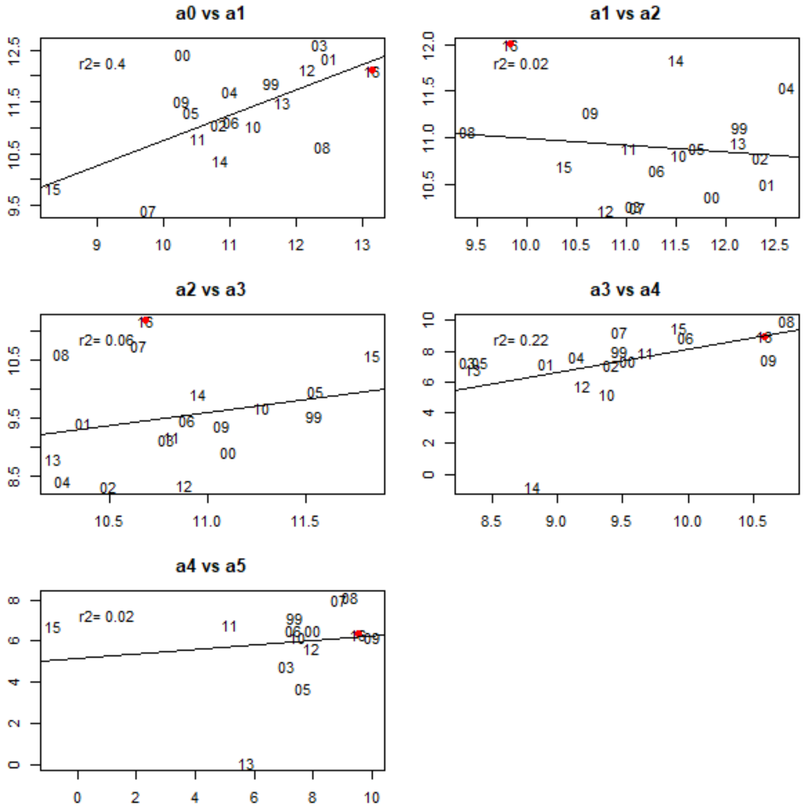


Figure 3.16b. Internal consistency of CPUE for survey model "Hard 24" (BITS-Q4 data).

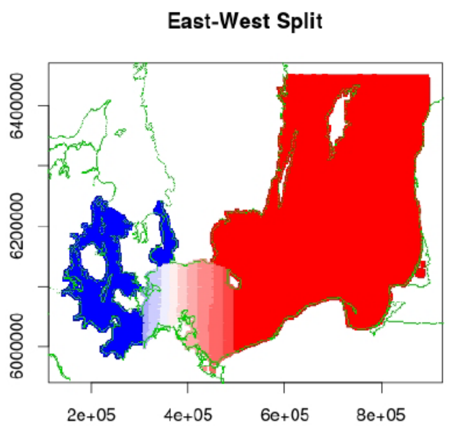


Figure 3.16c. Coverage area in the survey model "Hard 24".

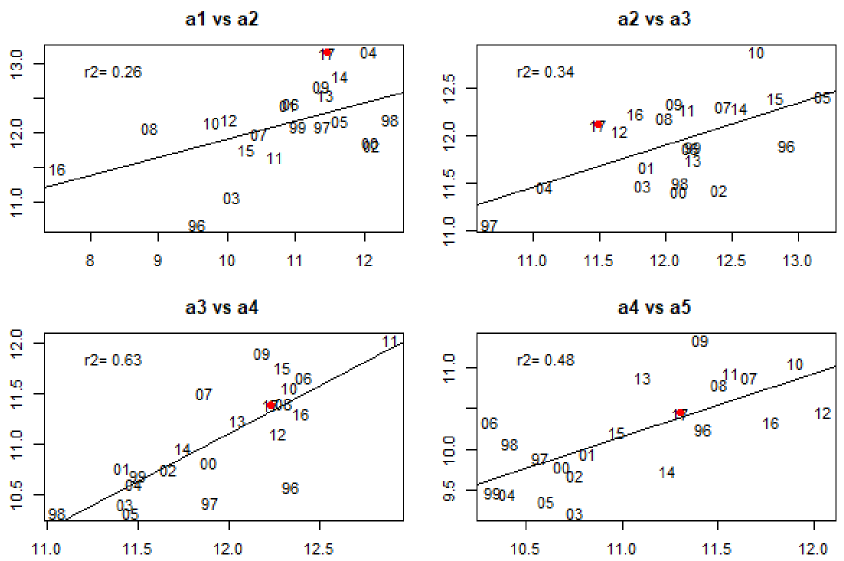


Figure 3.17a. Internal consistency of CPUE for survey model “Soft 24” (BITS-Q1 data).

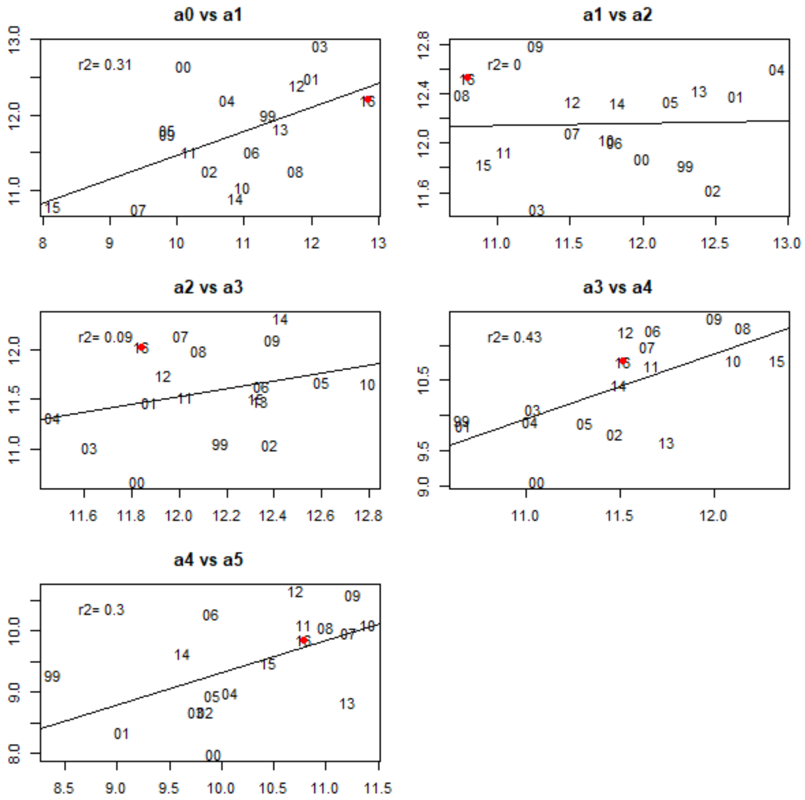


Figure 3.17b. Internal consistency of CPUE for survey model “Soft 24” (BITS-Q4 data).

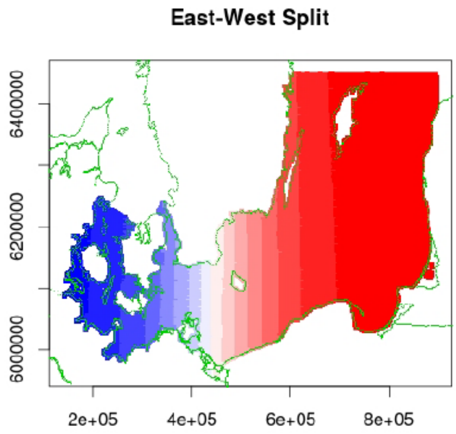


Figure 3.17c. Coverage area in the survey model “Soft 24”.

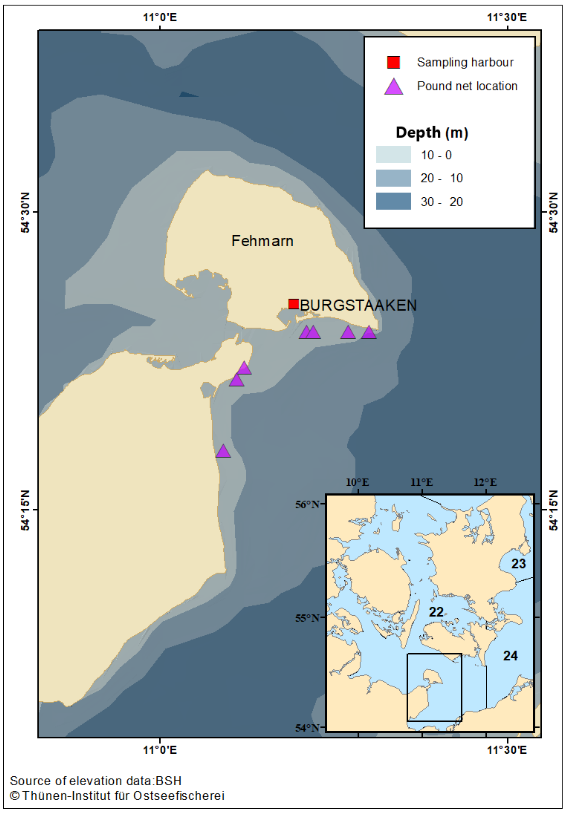


Figure 3.18. Location of poundnets off the coast of Fehmarn, from which samples were collected.

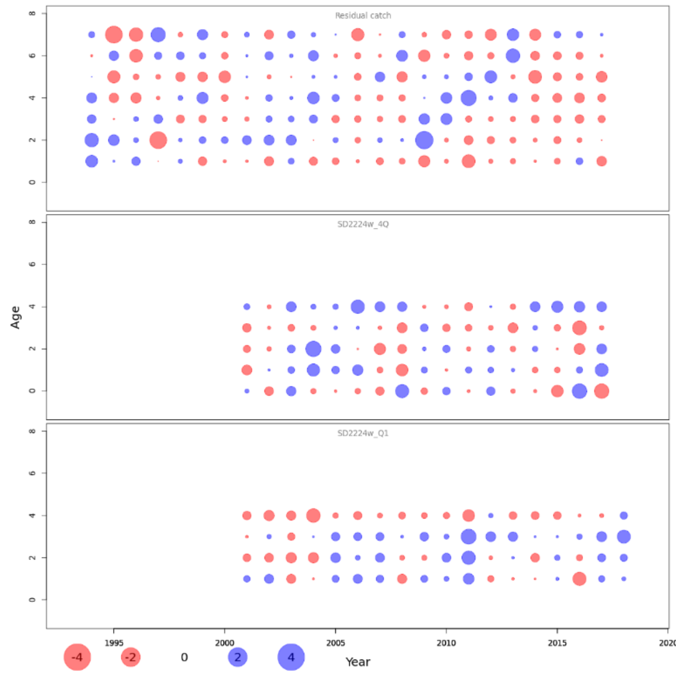


Figure 3.19. Western Baltic cod. Residual plot from the original assessment. A pattern is evident for the oldest age group in both the quarter 1. and quarter 4. survey. Further, the catch matrix has a pattern in the last 4 years.

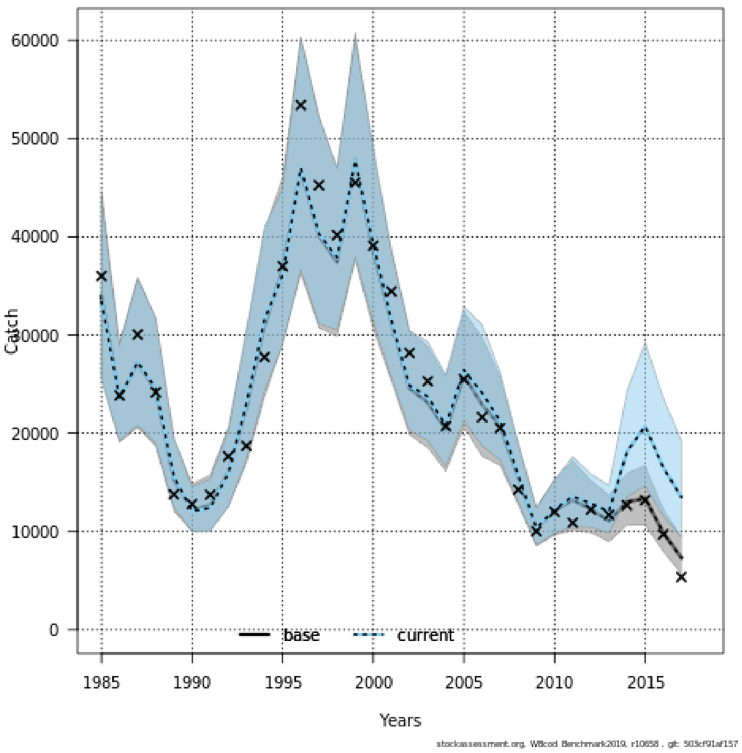


Figure 3.20. Western Baltic cod. Exploratory run with a catch multiplier for the years 2014–2017.The black line is the reported catches and the dashed blue line the estimated.

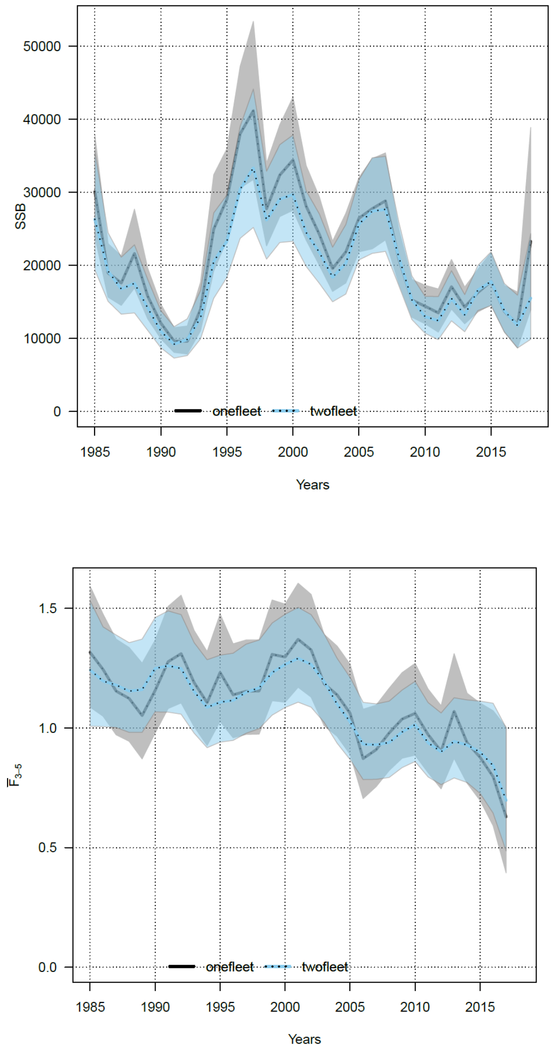


Figure 3.21. Western Baltic cod. Spawning-stock biomass and F (3-5) for the 2 fleet model (blue dotted line) compared to the one-fleet model (grey line).

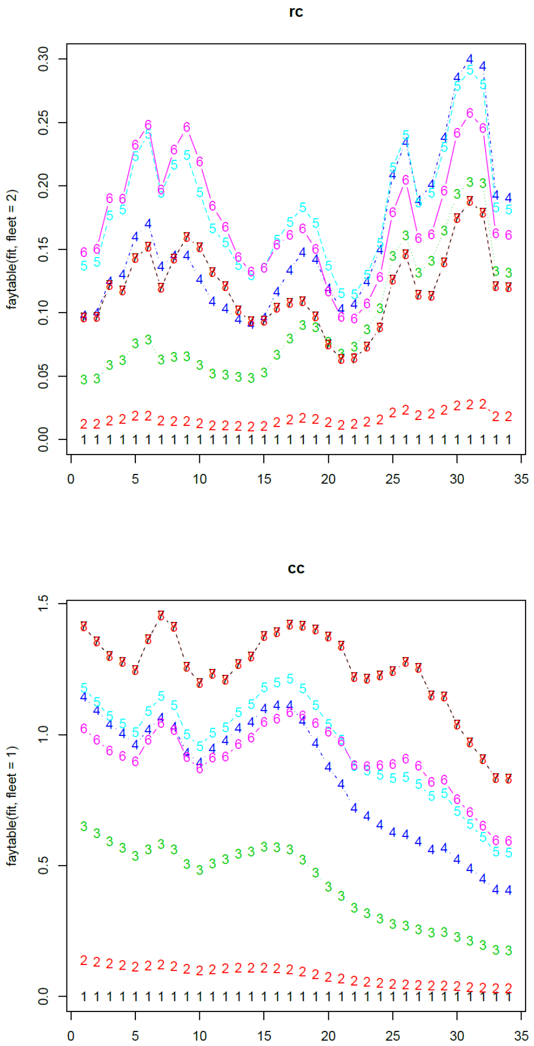


Figure 3.22. Western Baltic cod. F by age group in two-fleet model for recreational catch (left) and commercial catch (right).

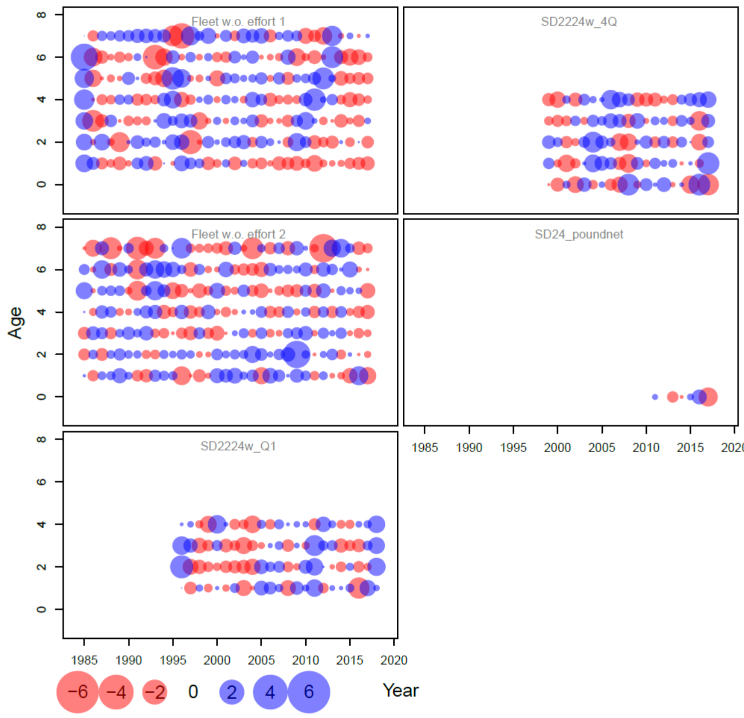


Figure 3.23. Western Baltic cod. Residuals from commercial catch matrix (effort 1), recreational catch matrix (effort 2), BITS-Q4 survey, BITS-Q1 survey and the poundnet survey from the final run.

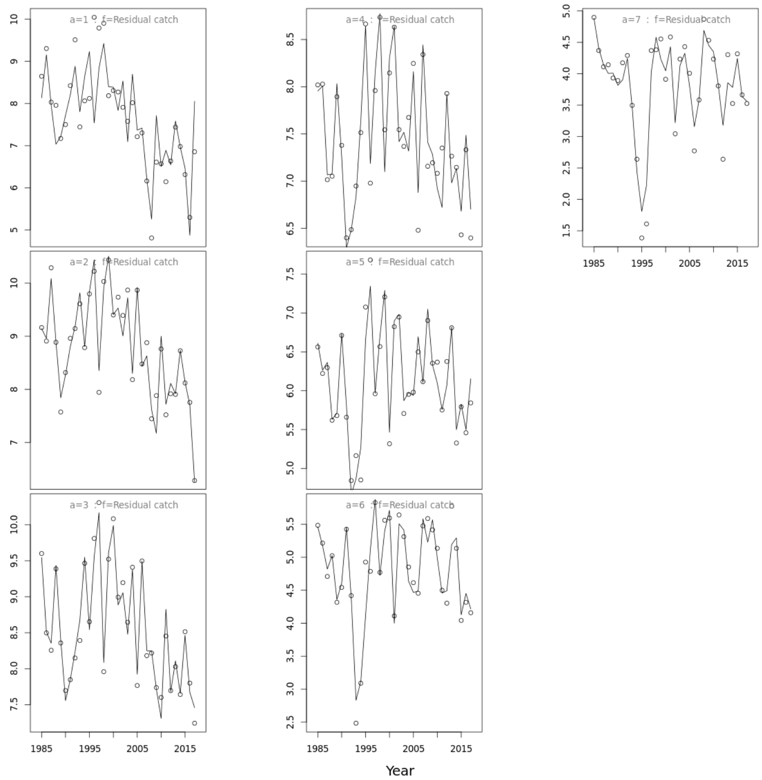


Figure 3.24. Western Baltic cod. Predicted line from catch matrix in relation to the observed dots (log scale)

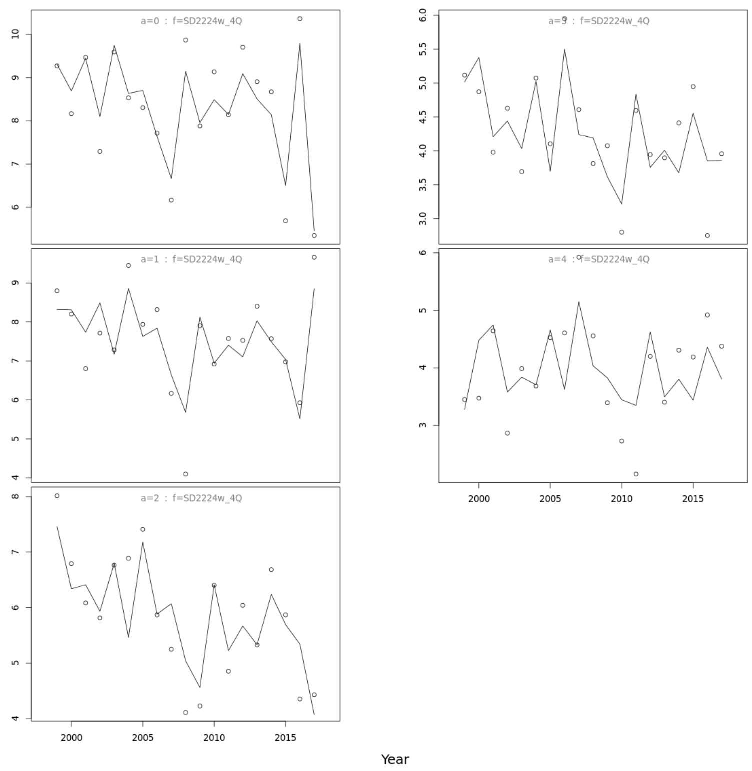


Figure 3.25. Western Baltic cod. Predicted line from BITS-Q4 survey in relation to the observed dots (log scale).

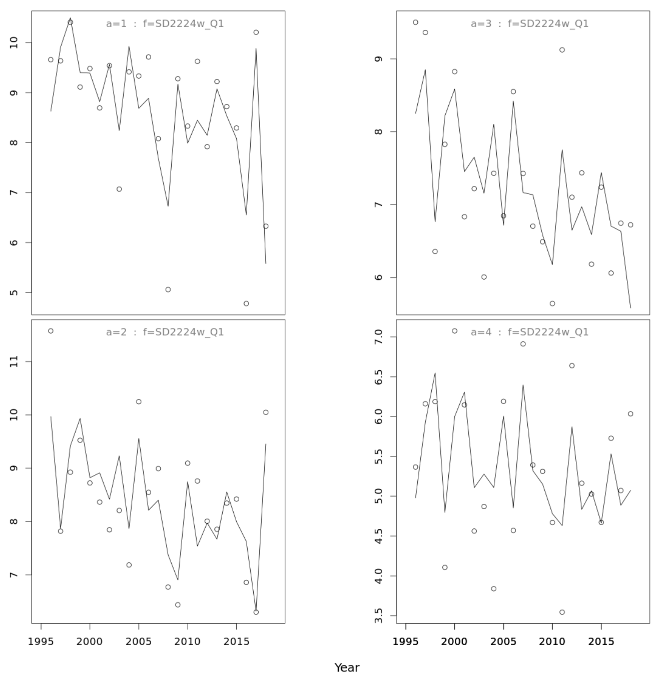


Figure 3.26. Western Baltic cod. Predicted line from BITS-Q1 survey in relation to the observed dots (log scale).

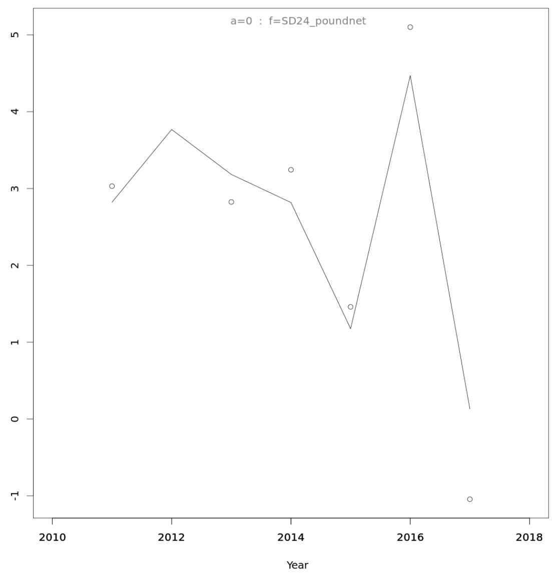


Figure 3.27. Western Baltic cod. Predicted line from poundnet survey in relation to the observed dots (log scale).

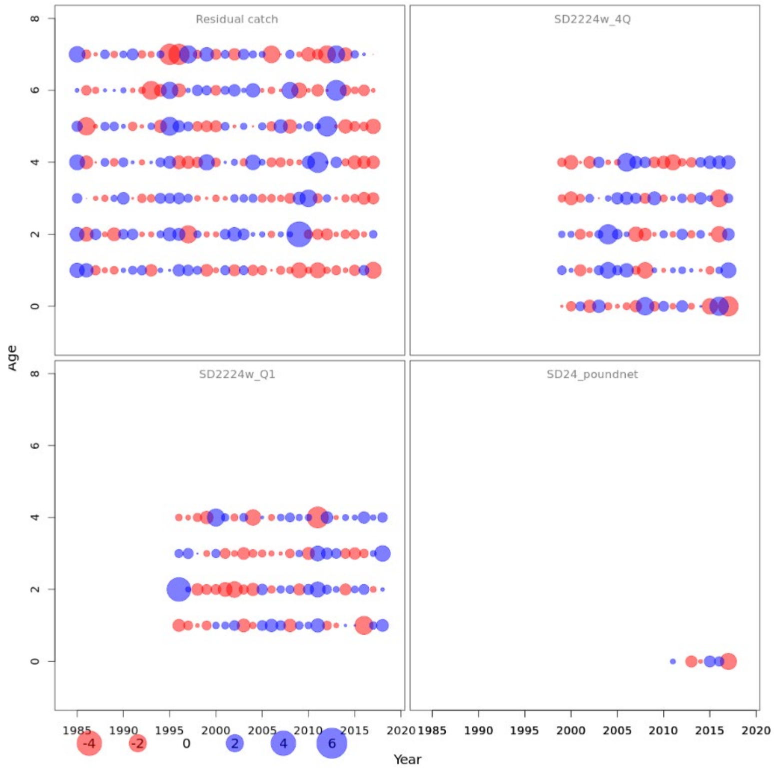


Figure 3.28. Western Baltic cod. Residuals from catch matrix, BITS-Q4 survey, BITS-Q1 survey and the poundnet survey from the final run.

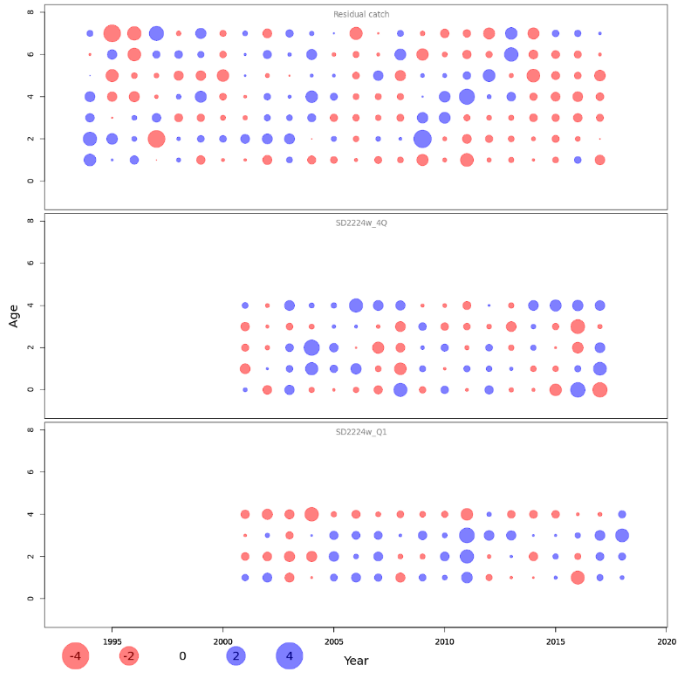


Figure 3.29. Residuals in survey and catch matrix with SPALY model settings

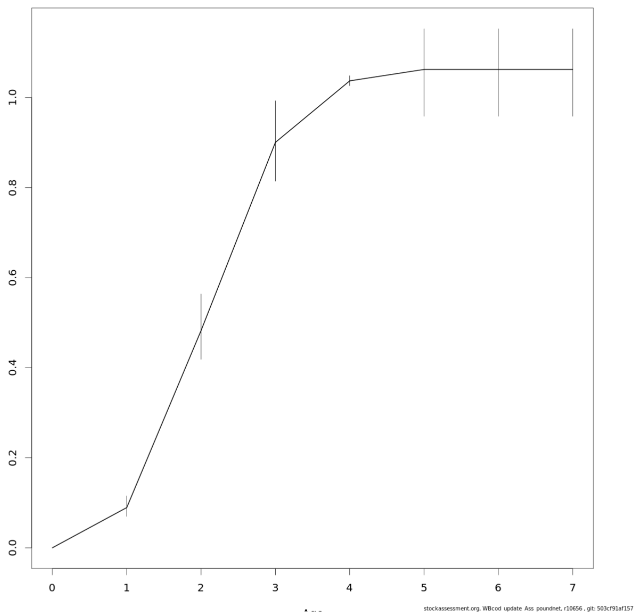


Figure 3.30. Western Baltic cod. Selection pattern from the combined fleet in the final run.

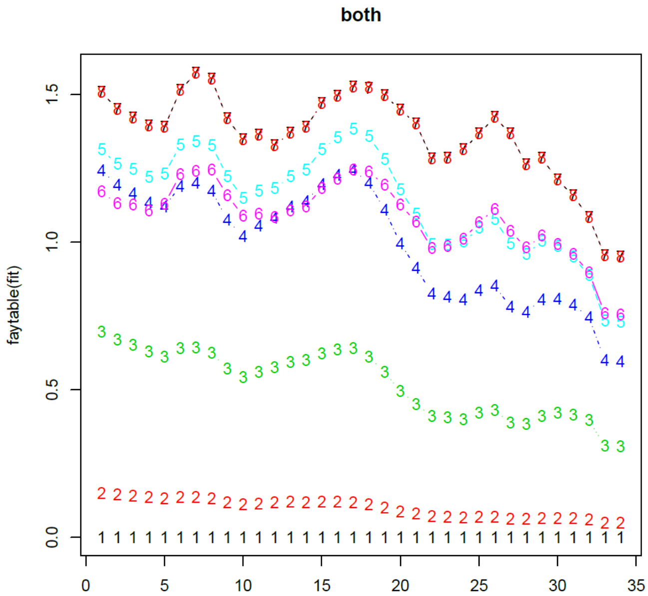


Figure 3.31. Western Baltic cod. Selection pattern from the combined fleet in the final run.

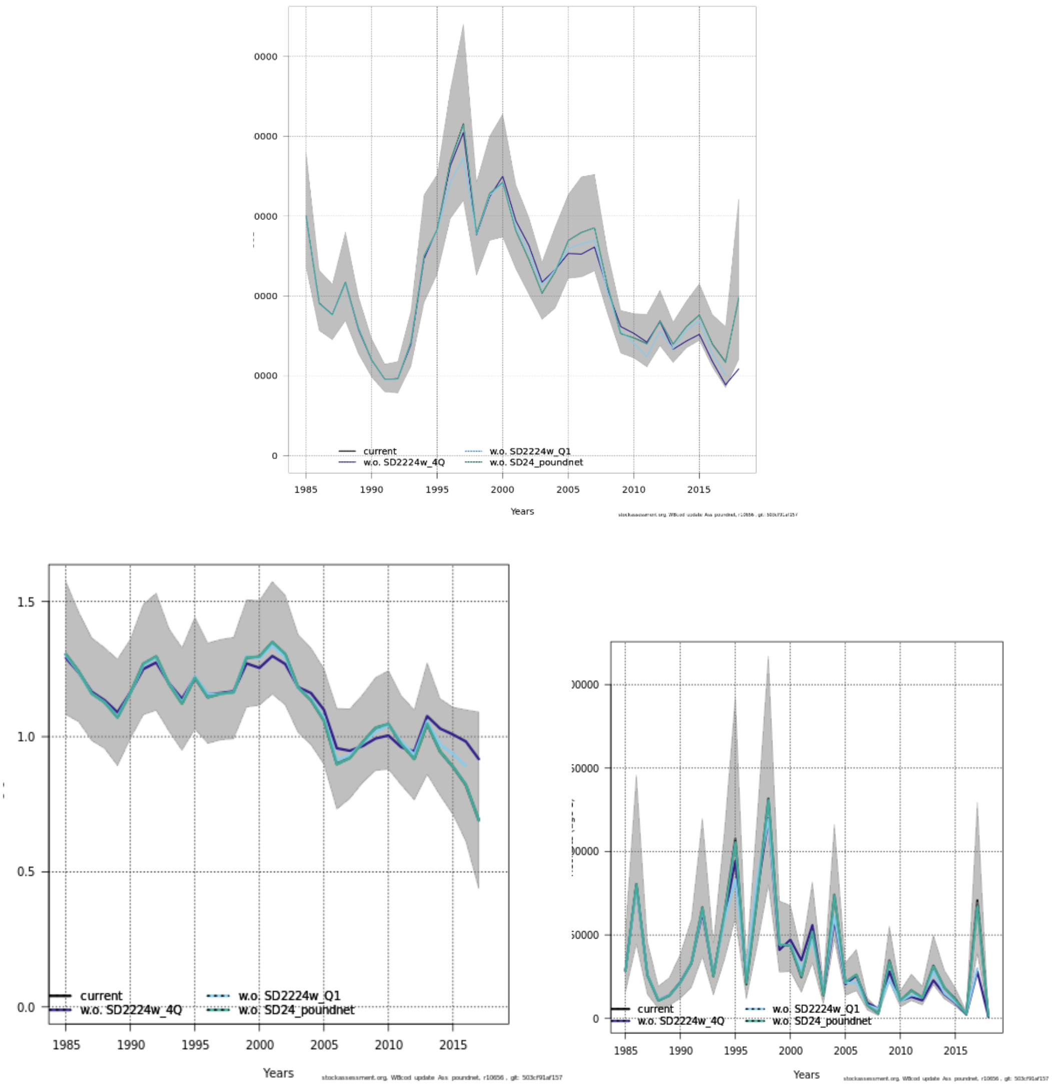
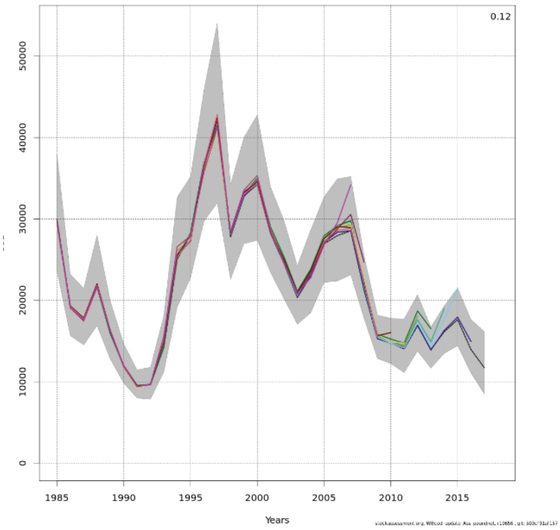
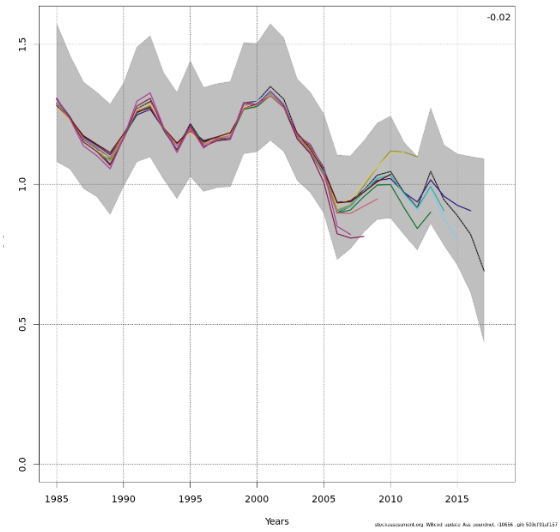


Figure 3.32. Western Baltic cod. Leave-one-out analyses (excluding one survey at a time) for spawning-stock biomass (SSB), F (3-5) and recruitment from the final run. The Q1 survey includes the latest years data point and therefore the leave one out (without this datapoint) cannot produce the last year.


0.12
50000
40000
30000
20000
10000
0
1985
1990
1995
2000
2005
2010
2015
Years


-0.02
1.5
1.0
0.5
0.0
1985
1990
1995
2000
2005
2010
2015
Years

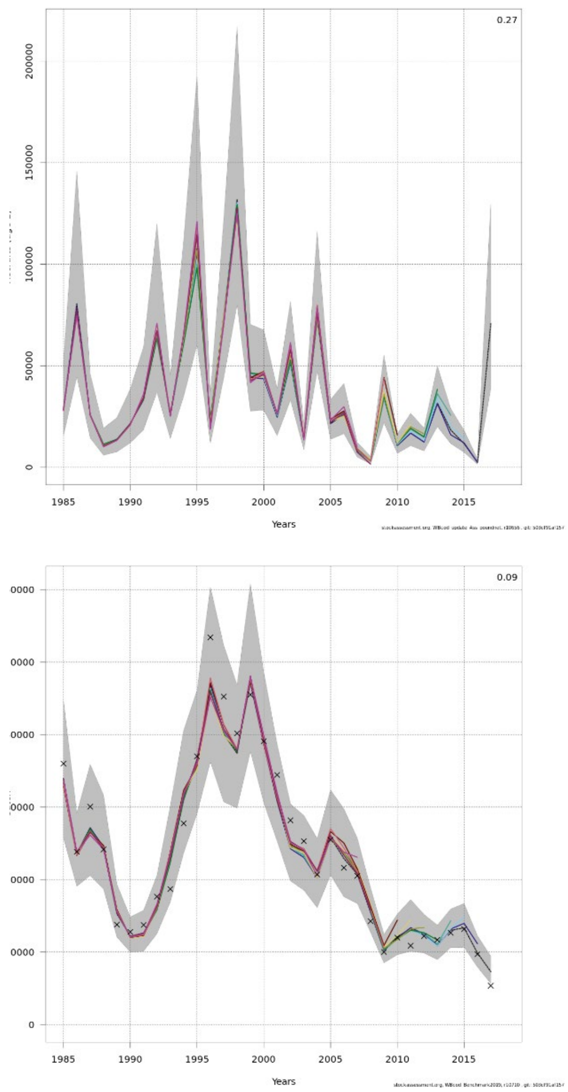


Figure 3.33. Western Baltic cod. Retrospective analyses for spawning-stock biomass (SSB), F (3-5), recruitment and catch from the final run.

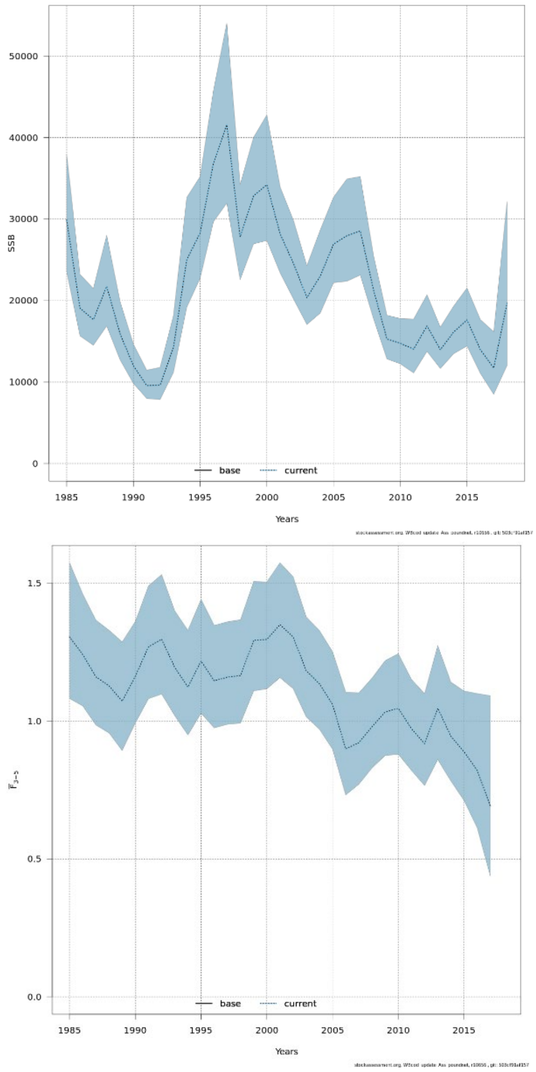


Figure 3.34. Western Baltic cod. Spawning-stock biomass (SSB) and F (3-5) from the final assessment.

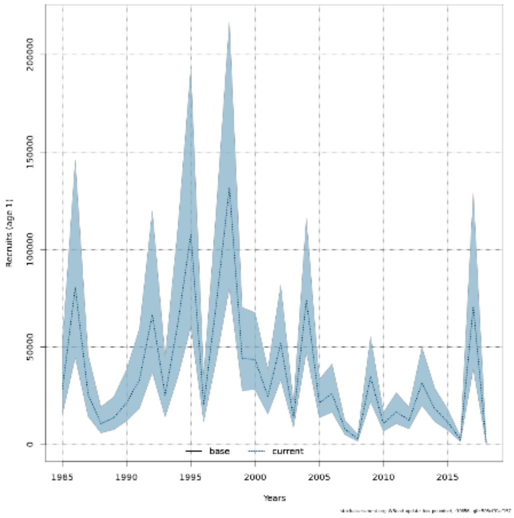


Figure 3.35. Western Baltic cod. Recruitment estimate (age1). For recruitment in the forecast years, it was decided to sample from the most recent 10 years recruitment estimate (1000 times with replacement)

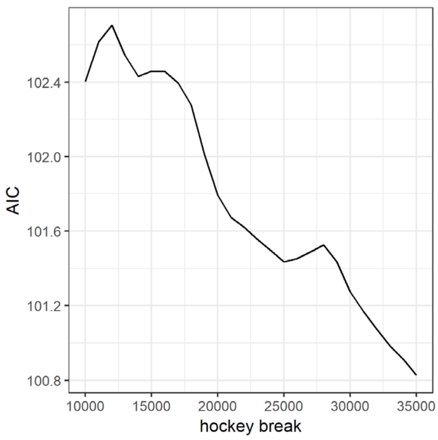


Figure 3.36. Western Baltic cod. Model to estimate hockey stick breakpoint

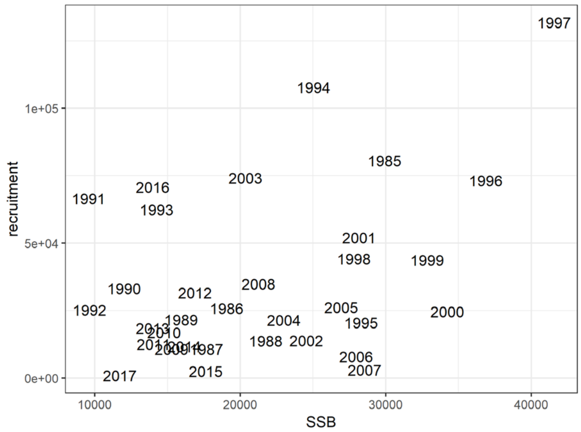


Figure 3.36. Western Baltic cod. Stock recruitment plot. The year classes 1991, 1993, 2016, and 2013 was used in an average to set B_{lim}.

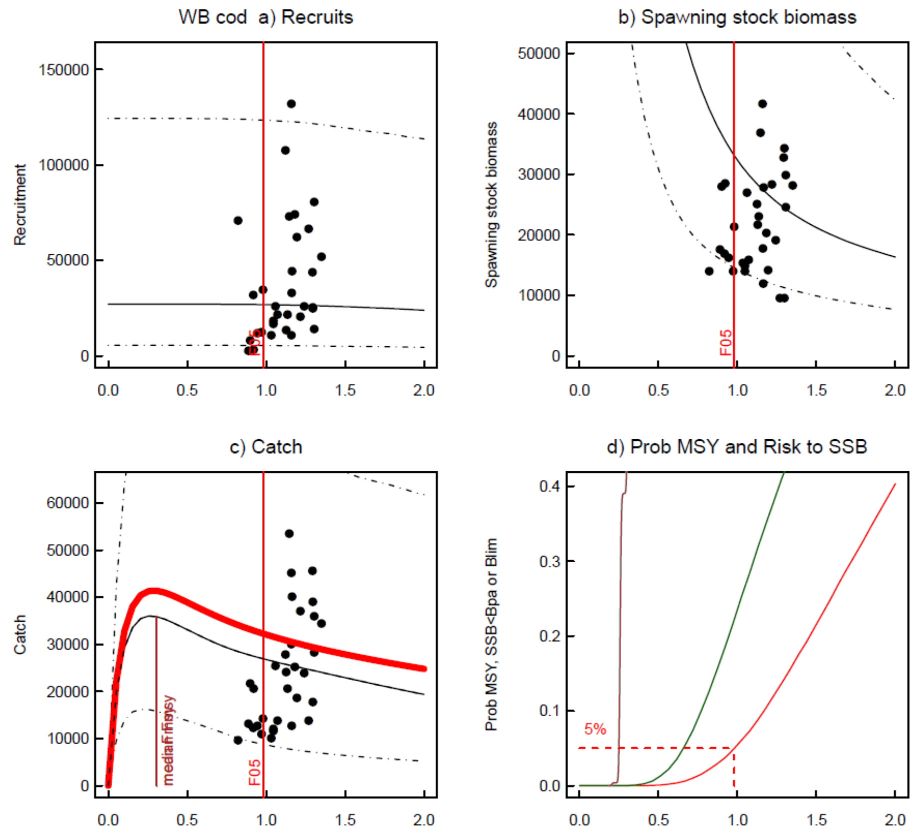


Figure 3.37. Western Baltic cod. Stock recruitment plot from EQSim. Data from 1985–2017

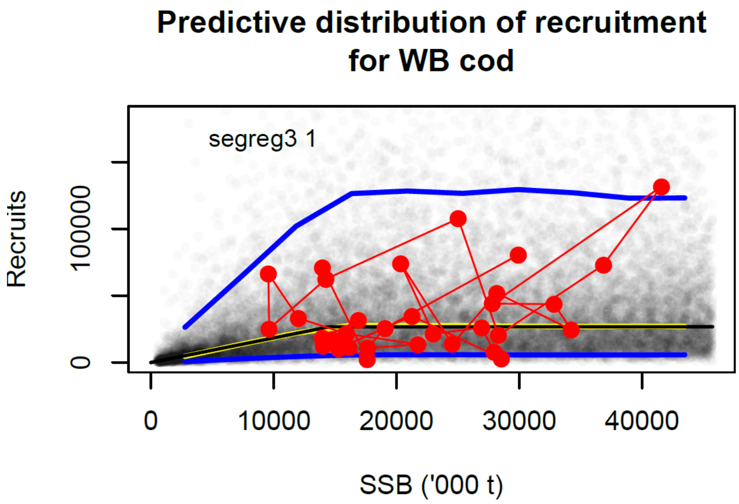


Figure 3.38. Western Baltic cod. Stock recruitment plot from EqSim. Data from 1985–2017

	COMB (DK+DE) split pct	
	pct_east (Area1)	pct_east (Area2)
1977	37	48.4
1978		48.5
1979		52.1
1980		
1981		60.3
1982		
1983		45.5
1984		
1985		43.8
1986		53.6
1987		
1988		
1989		
1990		
1991		
1992		45.8
1993		58.8
1994		53.5
1995		42.6
1996	34	51
1997		
1998	28	29
1999		
2000	29	51
2001		
2002		
2003		
2004		
2005	37.1	50.1
2006		
2007		
2008	54	80
2009		
2010	42.6	74.0
2011	49	85
2012		
2013	47	77
2014	49	75
2015	50	75.4
2016	42	76.6

Table 3.1. Years with split values combined from German and Danish split data.

	SD 22	SD23	SD24
CATON			
DK	1985–2008: Catch-per-year is calculated as the mean catch per year for the period 2009–2018, which is then weighted for each year with the number of Danish citizens being 18–65 years old.	1985–2008: Catch-per-year is calculated as the mean catch-per-year for the period 2009–2018, which is then weighted for each year with the number of Danish citizens being 18–65 years old.	1985–2008: Catch-per-year is calculated as the mean catch-per-year for the period 2009–2018, which is then weighted for each year with the number of Danish citizens being 18–65 years old.
	2009-2018: Statistics Denmark recall survey with adjusted estimates using correction factor from REKREA on-site studies on tour boats and private boats in SD23 in 2016–2018	2009–2018: Statistics Denmark recall survey with adjusted estimates using correction factor from REKREA on-site studies on tour boats and private boats in 2016–2018	2009–2018: Statistics Denmark recall survey with adjusted estimates using correction factor from REKREA on-site studies on tour boats and private boats in SD23 in 2016–2018
GE	1980–2004: reconstruction of the time-series is based on the average catch from 2009–2015. To account for the historic development (former GDR) catches in Mecklenburg-Western Pomerania were set to 20% from 1980–1991 with linear increase by 20% between 1991–1995		1980–2004: reconstruction of the time-series is based on the average catch from 2009–2015. To account for the historic development (former GDR) catches in Mecklenburg-Western Pomerania were set to 20% from 1980–1991 with linear increase by 20% between 1991–1995
	2005-2014: Annual catch is calculated on the basis of a mail-diary study (effort) corrected with annual license sales and using CPUE data from an annual on-site intercept survey		2005–2014: Annual catch is calculated on the basis of a mail-diary study (effort) corrected with annual license sales and using CPUE data from an annual on-site intercept survey
	2015–2017: Annual catch is calculated on the basis of a national telephone-diary study (effort) corrected with annual license sales and using CPUE data from an annual on-site intercept survey		2015–2017: Annual catch is calculated on the basis of a national telephone-diary study (effort) corrected with annual license sales and using CPUE data from an annual on-site intercept survey
SE		Tour boat sensus 2011–2018, Marina sampling of private boats 2017–2018	Marina sampling of private boats 2017–2018

Length			
DK	Same as German data	From on-site studies 2012, 2013, 2016, 2017, and 2018 used in combination with Danish and Swedish data. Face value data used 2012–2017. An average of the time-series was used on the historic data (1985–2012)	Same as German data

GE	1980–2004: pooled length distribution from 2005–2017 on-site measurement from DMAP national survey on-board tour boats, private boats (sea-based) and from self-sampling during fishing competitions (land-based)		1980–2004: pooled length distribution from 2005–2017 on-site measurement from DMAP national survey on-board tour boats, private boats (sea-based) and from self-sampling during fishing competitions (land-based)
	2005–2017: annual face values from on-site measurement from DMAP national survey on-board tour boats, private boats (sea-based) and from self-sampling during fishing competitions (land-based)		2005–2017: annual face values from on-site measurement from DMAP national survey on-board tour boats, private boats (sea-based) and from self-sampling during fishing competitions (land-based)
SE		Same as Danish data	
Age			
DK	Same as German data	Data from both Danish and Swedish recreational surveys, commercial landings and BITS survey. Data lacking from 1985–1990 and 2001–2003. Mean age length key based on the years 1991–1994 applied to the years 1985–1990. Mean age length key based on mean values on the years 1997–2000 and 2004–2008 applied to the years 2001–2003. Face value from 2016–2017.	Same as German data
SE		Same as DK	
GE	1980–2002: matching the recreational length distribution (total numbers-at-length) with ALK from BITS data for each year		1980–2002: matching the recreational length distribution (total numbers-at-length) with ALK from BITS data for each year
	2002–2017: matching the recreational length distribution (total numbers-at-length) with ALK from German commercial sampling data for each year		2002–2017: matching the recreational length distribution (total numbers-at-length) with ALK from German commercial sampling data for each year

Table 3.2. Overview of recreational data assumptions used by country and data source.

4 Eastern Baltic cod (SD 24-32)

4.1 Stock ID and substock structure

Eastern Baltic cod is genetically separated from Western Baltic cod. Eastern Baltic cod management area is SD25-32, while part of the population is also distributed in SD24, where both the Eastern and Western Baltic cod occur. Mixing of the two cod stocks in SD24 has been taken into account in ICES stock assessments since 2015. This was maintained at this benchmark, i.e. the both the catches and BITS survey indices included in the assessment include the fraction of Eastern Baltic cod caught in SD24. Specifics of the catch separation procedures are described under Western Baltic cod in Section 3, as well as outlined in Stock Annex.

4.2 Issue list

The aim of this benchmark was to establish analytical quantitative assessment (ICES Category 1) for the Eastern Baltic cod, which had been lacking since 2014. Therefore, all issues related to the input data needed for the assessment model were addressed. The main data issues relate to biological information on growth and natural mortality, as unclear developments in these processes have prevented analytical assessment in later years. Additionally, new tuning indices were developed for possible inclusion in the assessment model. Stock assessment models Stock Synthesis and SPICT were put forward for the benchmark as two possible model candidates. However, the benchmark later decided that Stock Synthesis should be the principal model to be evaluated while SPICT was considered as a backup model in case none of the Stock Synthesis model configurations would be considered appropriate. Moreover, reference points taking into account changes in productivity of the stock were also estimated as well as procedures to conduct short-term forecast.

4.3 Multispecies and mixed fisheries issues

Cod catches in SD24 are a mixture of Eastern Baltic and Western Baltic cod (see section 4.1). Eastern Baltic cod is affected by multiple species interactions (seals, benthic prey, sprat and herring). The ecological processes related to these interactions (further described in section 4.4) are considered to contribute to the currently low productivity of the stock (low growth, high natural mortality).

4.4 Ecosystem drivers

A number of changes in Eastern Baltic cod biology have been observed in later years, which include reduced nutritional condition of fish, maturation at a smaller size and increased parasite infestation due to grey seals. In addition, relative abundance of larger individuals in the population has sharply declined since 2012 (Eero *et al.*, 2015).

Nutritional condition of adult cod has been continuously declining since the early 1990s. However, since the mid-2000s, the proportion of cod with a very low condition index has rapidly increased (Eero *et al.*, 2012; Casini *et al.*, 2016). The decline in cod condition is evident in all offshore areas of the central Baltic. Hypothesized main reasons for deteriorating nutritional condition include:

(i) Low availability of fish prey in the main distribution area of cod, as sprat and herring are more northerly distributed with little overlap with cod (Eero *et al.*, 2012). (ii) Poor oxygen conditions that can affect cod growth directly via altering metabolism or via shortage of benthic prey (Casini *et al.*, 2016).(iii) Increased infestation with parasites, which is related to increased abundance of grey seals (Mehrdana *et al.*,2014; Howbowy *et al.*, 2016; Sokolova *et al.*, 2018).

Growth of Eastern Baltic cod is expected to have declined, associated with the above mentioned ecological processes, and additionally in relation to reduced size at maturation. The same factors have presumably contributed to an increase in natural mortality of the stock.

4.5 Stock assessment input data

4.5.1 Catch data

Total catch in tons in years 1946–1965 were obtained from Eero *et al.* 2007. From 1966 onwards, total catches in tons in SD25-32 are from earlier assessments (including discards and misreporting) (ICES WGBFAS 2018).The fraction of Eastern Baltic cod catch in SD 24 is added from 1965 onwards. The separation of the catches between Eastern Baltic and Western Baltic cod in SD24 is described in detail in the Western Baltic cod Section 3. For Eastern Baltic cod assessment, the assignment of landings taken in SD24 to the different stocks was extended back to 1965. For the historical period (1977–1995), proportions of Eastern Baltic and Western Baltic cod in SD24 are available from German historical survey (1977–1986), supplemented by stock proportions derived from BITS survey (1992–1995). The extrapolation back to 1965 applies average proportions of landings in subareas 1 and 2 in SD24 (Figure 3.1), and average stock mixing proportions from 1977–1979. Before 1965, cod landings in SD24 were not available. However, the fraction of SD24 in Eastern Baltic cod landings is generally low before the 1990s (<5% in most years) (Figure 4.1). Therefore, no catches of Eastern Baltic cod in SD24 prior to 1965 were assumed.

Catch by Fleets:

Total catch biomass is divided between Active (trawls) and Passive (mainly gillnets) fleets from 1987 onwards. The gillnet fishery for cod in the Baltic Sea developed mainly in the 1990s. Thus, although some fraction of landings was taken by gillnets also in earlier years, their share in total catches was small and therefore the catches of the passive gears before 1987 were assumed to be zero.

Years	Description
1946–1986	All catches allocated to Active fleet
1987–1992	The fraction of Active/Passive in total landings in each of these years is based on Danish data (Maris Plikshs, pers. comm). This is used to divide total catch to the two fleets.
1993–1997	The total catch of the stock is allocated to Active/Passive, based on average contribution of these fleets to the total catches in years 1992 and 1998-2000.
1998–2017	Total catch of the stock is allocated to Active/Passive, based on the relative share of the gears in catches in a given year, reported to WGBFAS.

Catch by Quarter:

Years	Description
1946–1973	The average quarterly distribution of total catch observed in 1974–1978 was applied for all the years in 1946–1973.
1974–1999	The total catch is allocated to quarters based on quarterly distribution of the catch in the Baltic multi-species modelling (WGSAM) dataset, which is quarterly resolved. These data are not distinguishing between Active and Passive, thus the same average quarterly distribution of total catch is applied both for Active and Passive fleet.
2000–2017	The total catch is divided between quarters based on quarterly distribution of the catch provided to WGBFAS, separately for Active and Passive.

4.5.2 Age and length composition of catch

Age composition of catches is included in the model for 1946–2006 (effectively until 1999 as the age composition of catches for 2000–2006 is set to not contribute to the model likelihood and are treated as “ghost fleet” by Stock Synthesis). Age composition of the catches for later years is not included due to increased discrepancies between different countries’ age readings, which were identified to have occurred after 2007. Age reading has always been challenging for Eastern Baltic cod. However, this was not found to cause major biases in assessments in former times (Reeves, 2003) and ICES WKIDEBCA (2018) also concluded that catch-at-age data prior to 2007 are reasonable to be used in the assessment.

The data on the age structure of the catches are for the area SD 25-32. The same age structure of the catches as in SD 25-32, by year, is assumed to apply also for the Eastern Baltic cod catch component in SD 24. As a first step, age structure of the total annual catch was compiled. Differently from previous stock assessments that started from age 2, data for age 1 were compiled as well, as Stock Synthesis model starts from age 0.

Age 1:

Years	Description
1946–1973	Based on historical landings at age data compiled from national reports and literature (Eero *et al.,* 2007)
1974–2006	Catch-at-age from multispecies assessment input (WGSAM)

Age 2+:

Years	Description
1946–1965	Historical landings at age data compiled from national reports and literature (Eero *et al.,* 2007)
1966–2006	Catch-at-age used in former stock assessments (WGBFAS)

Age compositions by fleet:

Years	Description
2002–2006	Total catch of a given age was divided to Active and Passive, based on fleet specific catch-at-age information provided in WGBFAS reports.
1998–2001	Landings-at-age: extracted from WGBFAS reports, separately for Active and Passive. Discards-at-age: Available only for fleets combined. These total discards for each age were distributed between Active and Passive based on total numbers of discards in Active and Passive fleets, assuming the same age distribution of discards in both Active and Passive gears.
1993–1997	Average relative contribution of Active and Passive fleets to total catch of each age group in 1998–2005 was applied to distribute total catch-at-age of a given age group to fleets, as the relative contribution of fleets to the total catch was stable in 1993–2005.
1987–1992	The age structure in both Active and Passive was set to be the same as in the total catch-at-age for these years.

Age composition by quarter and fleet:

Years	Description
1974–2006	Total catch of each age group (by fleet) was divided between quarters based on multispecies input data, which is quarter specific. These annual proportions of quarters were applied to distribute both Active and Passive catch-at-age to quarters. Thus, quarterly proportions were specific to each age group and year, but the same for Active and Passive.
1946–1973	Average age specific quarterly proportions recorded for 1974–1978 were applied for catch-at-age in all years in this period.

Length compositions:

Data on length compositions of the catches in SD 25-32 are available from 2000 onwards, by Active and Passive fleet and by Quarter. The national data are uploaded in Intercatch database (IC). The landings that have not been specified in IC whether active or passive were all allocated to Active. The Eastern Baltic cod catches in SD 24 are assumed to have the same length distribution as in SD 25.

4.5.3 Conditional age at length (age–length key)

Age length keys are used in Stock Synthesis model from 1991 onwards to inform the estimated deviations in von Bertalanffy growth parameters. The ALKs used in the reference run are based on age readings from BITS surveys, available in DATRAS (ALK3 in WD1). Both ALKs from Q1 (1991–2017) and Q4 (1998–2017) were included.

4.5.4 Surveys

The tuning series included in Stock Synthesis model are described in the table below. For #BITSQ1 and #BITSQ4, length composition data are included as well. The time-series of tuning indices that cover the period up to the present are shown in Figure 4.2.

Fleet name	Years	Description
#BITSQ1	1991–2018	Baltic International Bottom Trawl Survey, Q1, data for SD 25-32, including the area east of 13 degrees latitude in SD 24. Modelled indices of total abundance. Method for survey modelling is described in WD2.
#BITSQ4	1993–2017	Baltic International Bottom Trawl Survey, Q4, data for SD 25-32, including the area east of 13 degrees latitude in SD 24. Modelled indices of total abundance. Method for survey modelling is described in WD2.
#TrawlSurvey1	1975–1992	CPUE (kg*h–1) by German RV Solea in SD 25 (Thurow and Weber, 1992)
#TrawlSurvey2	1978–1990	CPUE (g/hour) from bottom trawl surveys by the Swedish Board of Fisheries and Baltic Fisheries Research institute (BaltNIIRH), SDs 25–28, yearly average. The index refers to total cpue in biomass of all length groups caught in the survey (Orio *et al.,* 2017).
#CommCpue1	1948–1956	Commercial CPUE (kg/h) of former USSR , February–June (Dementjeva, 1959)
#CommCpue2	1957–1964	Commercial CPUE (kg/h) of former USSR in Gdańsk area, February-June (Birjukov, 1970)
#CommCpue3	1954–1989	Commercial CPUE (kg/day) of USSR (Latvian republic), SDs 26-28, annual average (Lablaika *et al.,* 1991)
#SSBEggProd	1986–2017	SSB indices based on annual egg production method. Used in SS model to represent spawning-stock biomass trends (survey type 30 in SS). Data from ichthyoplankton surveys. Calculation of SSB indices described in WD3.
#Larvae	1987–2017	Abundance of larvae during peak spawning, used in SS as prerecruit survey (survey type 32). Data from ichthyoplankton surveys. Calculation of the index is described in WD4.

4.6 Stock Assessment: Stock Synthesis

4.6.1 Model configuration and assumptions

General model specifications

The assessment of the Eastern Baltic cod (SD24-32) was conducted using the Stock Synthesis (SS) model (Methot and Wetzel, 2013). Stock Synthesis is programmed in the ADMB C++ software and searches for the set of parameter values that maximize the goodness-of-fit, then calculates the variance of these parameters using inverse Hessian and MCMC methods. The assessment was conducted using the 3.30 version of the Stock Synthesis software under the windows platform.

The Stock Synthesis model of Eastern Baltic cod is a one area quarterly model where the population is comprised of 15+ age classes with both sexes combined. The model is a length-based model where the numbers at length in the fisheries and survey data are converted into ages using the von Bertalanffy growth curve. The model is run in quarterly steps to account for the growth of individual cod throughout the year. Although the quarterly model assumes the same length-selectivity in the fishery throughout the entire year, the derived age-selectivity changes by quarter as the cod grows, which is something a yearly model cannot account for. SS assumes multinomial likelihoods for the proportions-at-length in catches and survey data.

The last age class (i.e. 15+) represents a "plus group" in which mortality and other characteristics are assumed to be constant. The model starts in 1946 and the initial population age structure was assumed to be not in an unexploited, equilibrium state so that the initial fishing mortality was assumed to be 0.6 in the model based on Eero *et al.* (2008). Initial catches (i.e. catches before 1946) were assumed as the average of the 10 preceding years (1936–1945) based on Eero *et al.* (2008). Fishing mortality was modelled using the hybrid method that the harvest rate using the Pope's approximation then converts it to an approximation of the corresponding F (Methot and Wetzel, 2013).

Spawning stock and recruitment

Spawning-stock biomass is estimated for spawning time (month 5 is used as an average for the entire period). Sex ratio is set to 50% females and males. Recruitment was derived from a Beverton and Holt (BH) stock recruitment relationship (SRR) and variation in recruitment was estimated as deviations from the SRR. Recruitment deviates were estimated for 1946 to 2016 (71 annual deviations), representing the period for which age and length compositions are available. Recruitment deviates were assumed to have a standard deviation (σR which corresponds to the stochastic recruitment process error) of 0.6. For the period 1935–1945, recruitment was derived directly from the SRR and from the initial catches and fishing mortality. The reference model assumed a level of steepness (h) of 0.99 for the SRR, assuming that recruitment is mainly environmentally driven in EBC. Settlement time for recruitment in the reference model is set to month 8 as an average for the entire period (3 months after the spawning time).

Growth

Growth parameters were fixed for the period 1946–1990, at values estimated using historical tagging data (see WD5 for details) (Table 4.2). The tagging estimates covered the period 1955-1970 (L_{inf} = 125.27, k= 0.10). Deviations in both L_{inf} and k were estimated between 1991 and 2018 when age–length keys were available from BITS surveys. Age–length Key (ALK) therefore is used to inform the estimation of growth deviations from 1991 onwards. Numbers of fish in ALK are used as sample size for each year. The variance in length-at-age was fixed for older fish and estimated for younger individuals (Table 4.2). Length at minimum age (A_{min}) was first estimated in Stock Synthesis model, and then fixed at the estimated values (Table 4.2).

The parameters a and b in length-weight relationships are estimated from Q1 BITS survey, pooled for SD 25-32. The parameters were estimated for each year, after which the data were averaged by 3-year blocks. These externally estimated parameters were used as inputs in the model.

Natural mortality

Natural mortality was assumed to be age dependent and it was estimated using methods described in Then *et al.* (2015) and Lorenzen (1996) for the historical period (1946–1999) (see WD6 for details). Then *et al.* (2015) estimation of M is based on maximum age (*tmax*) and parameters of the von Bertalanffy growth curve. The Lorenzen type (Lorenzen, 1996) of M-at-age function assumes a declining relationship between M and the mean weight of fish in successively older age classes. Natural mortality of the reference model was assumed to be equal to the average of the two methods (*tmax* and *growth*) scaled using Lorenzen (1996) (Figure 4.3**)**. For the reference run of Stock Synthesis, age breakpoints 0.5, 1.5, 5.5, and 15.5 were used. Natural mortality from 2000 to 2018 for-age break 5.5. was estimated within the model as annual deviations from the historical values. For the other age-breaks, M is kept constant for the entire time-series (Table 4.2).

Maturity

The input for maturity is L_{50} (length at 50% mature) and the slope of the maturity ogive curve. These are estimated outside the stock assessment model from BITS Q1 data, for females and males combined. L_{50} of Eastern Baltic cod has substantially declined over time, which is captured by using time blocks in the assessment model (Table 4.2). For the slope, a constant value (0.23) is used for the entire period. The change in L_{50} estimated from BITS Q1 was validated with data from German CoBalt survey, conducted closer to the spawning time. These results also confirmed the decline in L_{50} observed in BITS.

Selectivity

Fishery selectivity is assumed to be length-specific and time-invariant. For both the trawlers (i.e. active gears) and the gillnetters (i.e. passive gears) selectivity was estimated assuming a logistic function that constrains the older age classes to be fully selected ("flat top"). A logistic selectivity was also used for BITS surveys (both quarter 1 and quarter 4). Selectivity of Trawlsurveys 1 and 2 was assumed to mirror selectivity of BITS Q1 survey, while selectivity for commercial CPUE1, 2 and 3 was assumed to mirror selectivity of the active gears.

4.6.2 Uncertainty measures and likelihood

The total likelihood of the model is composed of a number of components, including the fit to the survey and CPUE indices, fishery length frequency data, age compositions and catch data. There are also contributions to the total likelihood from the recruitment deviates, priors on the individual model parameters and tags (if any). The model is configured to fit the catch almost exactly so the catch component of the likelihood is very small. Details of the formulation of the individual components of the likelihood are provided in Methot and Wetzel (2013).

The CV of catch was set to 0.05 for all years. No meaningful information is available on the annual sample size associated with age or length distribution data for commercial catches. Therefore, in Stock Synthesis, the same value (100) is applied for each quarter and fleet in all years.

The average CV of the BITS survey indices was assumed to be equal to 0.15 while the yearly deviation of the coefficient of variation of the BITS survey indices was estimated as part of the modelling of the survey indices outside the stock assessment model (WD2). Numbers of hauls in BITS in each year were used as input for sample size associated with BITS length distribution data.

For the remaining surveys and CPUE indices, the CV was estimated internally for the reference model, except for the larvae index, for which the CV was set to 0.3. Data weighting is an important component of integrated stock assessment models. The weighting method used for the size-composition data followed the advice of Francis (2011) (Method TA1.8). For weighting the conditional age-at-length data we used the Francis-B approach described in Punt (2017). Iterative application of model fitting and reweighting occurred three times to explore the effects on successive estimates of the data weighting coefficient for each composition dataset. Weights from the second iteration were used for the results reported here because this iteration resulted in the smallest gradient for the objective function to be minimized among the three iterations of the model. The Hessian matrix computed at the mode of the posterior distribution was used to obtain estimates of the covariance matrix, which was used in combination with the Delta method to compute approximate confidence intervals for parameters of interest.

4.6.3 Exploratory runs

4.6.3.1 Exploratory runs with tuning indices

In former assessments, BITS survey indices have been obtained from DATRAS database. At benchmark 2019, modelled indices were introduced instead. There are some differences in length distributions between the modelled and DATRAS indices in some years (Figure 4.4), however this has only a minor influence on the assessment results (Figure 4.5). Also, including the fraction of cod from SD24 (i.e. fish caught east from 13 degrees longitude) in survey indices has a little impact on assessment results, compared to just using the indices for SD 25-32 (Figure 4.5). In the final assessment, the index covering SD25-32 and the area east from 13 degrees longitude in SD24 is used, which is consistent with the indices used for the Western Baltic cod stock.

The runs leaving out one series of the tuning indices at a time, did not reveal conflicting information between the different survey indices (Figure 4.6).

4.6.3.2 Exploratory runs with different values of natural mortality

The following runs with different M values and configurations were explored:

Model	Age-breaks	M value/method	DevAge	DevPeriod
Reference	0.5,1.5,5.5,15.5	average Tmax/growth	5.5	2000-2018
M02	0.5,1.5,5.5,15.5	0.2 at age 5.5	5.5	2000-2018
Mtmax	0.5,1.5,5.5,15.5	tmax method	5.5	2000-2018
Mgrowth	0.5,1.5,5.5,15.5	growth method	5.5	2000-2018
M02a25	0.5,1.5,2.5,15.5	0.2 at age 2.5	2.5	2000-2018
Ma35	0.5,1.5,3.5,15.5	average Tmax/growth	3.5	2000-2018
Ma45	0.5,1.5,4.5,15.5	average Tmax/growth	4.5	2000-2018
Ma65	0.5,1.5,6.5,15.5	average Tmax/growth	6.5	2000-2018
Mdev05	0.5,1.5,5.5,15.5	average Tmax/growth	5.5	2005-2018
MdevAgeBoth	0.5,1.5,5.5,15.5	average Tmax/growth	5.5, 15.5	2000-2018

These alternative runs generally did not improve the likelihood or convergence of the model, or the fits to different data sources, compared to the reference run. Conversely, the models with lower M values (M02a25, Mgrowth) had much worse convergence than the reference run. Thus, assuming M at 0.2 for all ages 2+ as used in earlier stock assessments, resulted in very poor convergence of the model. The estimated deviations in M and growth parameters were relatively similar in these alternative runs (Figure 4.7). The differences in the level of M have an impact on the resulting stock estimates, with higher stock at a higher M, as expected. Especially the absolute level of recruitment is estimated much higher with a higher M (Figure 4.8). However, the level of historical M mostly acts as a scaling factor for the stock estimates, with the F (and SSB) dynamics being less sensitive to the historical M level (Figure 4.8). The M in the reference run is based on average of the two type of methodologies suggested in the literature for estimating M (i.e. tmax and growth method, WD6). Thus, the average of the two methods was considered as a reasonable assumption for historical M (1946–1999).

Estimating deviations in M also for the oldest age (15.5) resulted in very little difference, as the deviations from the historical value were estimated to be very small. The age-break 5.5 is chosen somewhat arbitrarily and other breaks could be used instead. However, the alternative runs with other age breaks showed little impact on the model results.

During the benchmark meeting, an additional run was made adding an additional age breakpoint (0.5, 1.5, 3.5, 6.5, 15.5) and estimating deviations both for 3.5 and 6.5 since 2000. This did not improve the model and the results in terms of stock trends were very similar to the reference run. Therefore, it was decided to keep the M values and configuration as in the reference run for the final Stock Synthesis assessment.

4.6.3.3 Exploratory runs with growth – age length key

Age information from traditional age readings is considered uncertain, especially in later years. Therefore, different options for ALK were explored, and robustness of the assessment results to the uncertainties in age information was evaluated. Briefly, ALK1 is not using age readings after 2006, but this ALK is constructed based on expected changes in growth due to observed changes on biology of the stock and environmental conditions. ALK2 is based on age readings of selected countries. ALK3 (used in reference run) is based on all BITS age readings available in DATRAS database. Details on these ALKs are provided in WD1.

Another issue that was explored was whether changes in growth should be estimated for all years up to present, or only until around 2012 and kept constant thereafter, because very few larger individuals are available in the ALK (and in the stock) for later years, with increasing difficulties to estimate growth parameters.

Preliminary estimates of growth parameters in recent years (Linf~ 80 cm, K~ 0.15) from the tagging program (TABACOD project) were also available, which were used to validate the estimated change in growth. However, the present growth parameters estimated from tagging are considered to be an overestimate of the Eastern Baltic cod growth. This is because the tagged fish most likely include individuals of Western Baltic cod. Even if it is few individuals, it can have large impact on the von Bertalanffy growth estimates, as the sample size is relatively small, especially for larger cod. Nevertheless, the recent tagging confirms the decline in growth compared to the estimates from historical tagging. The growth estimates from tagging will be improved in future, when genetic analyses of the returned tagged cod will be conducted. For this reason, the growth estimates from recent tagging were currently not included in the reference model, but only in an exploratory run.

The following exploratory runs were conducted:

Model	Basis for age info	Period of ALK used	Deviations estimated
Reference	BITS age readings	1991–2018	1991–2018
ALK1ref	constructed, no age data used after 2006	1991–2018	1991–2018
ALK2ref	Age reading of selected countries	1991–2018	1991–2018
ALK3devStop12	BITS age readings	1991–2018	1991–2012
ALK3NoDat12	BITS age readings	1991–2012	1991–2012
ALK3dev17dat12	BITS age readings	1991–2012	1991–2018
VBLTag	BITS age readings +tagging	1991–2010 ALK , 2011–2018 tagging	1991–2018

Exploring the basis for age/growth information

The run with ALK1 had very poor convergence. The difference in ALK1 compared to the other options of ALK is that it is smoothed (while other ALKs are based on raw data), and the same ALK is inserted for multiple years, not accounting for year-class effects. Thus, ALK1 would not be most appropriate to be used in the final model, but was explored for comparison, as the advantage of ALK1 is that it is not using age readings for later years. The estimates of M and L_{inf} from ALK1 were similar to the reference run, with somewhat higher k value in ALK1 run (Figure 4.9). ALK2 estimated somewhat lower L_{inf} and higher M compared to reference model, but the differences where not substantial (Figure 4.9). The von Bertalanffy growth parameters from tagging for the latest years show higher L_{inf} and lower k compared to the other runs. However, as explained earlier above, the present growth estimates from the tagging are preliminary and likely overestimate L_{inf}, and underestimate k (the two being negatively correlated). Nevertheless, also the growth estimates from TABACOD tagging show a decline in growth, in line with the other runs with ALKs (Figure 4.9). The runs with higher growth generally estimate higher M (which is expected as the two parameters are confounded). It is recognized that the exact values for growth and M estimated from the model, i.e. separating between the two, is associated with some uncertainty, as the ALK information to inform growth is imprecise and the two parameters are biologically interlinked and thus confounded in the model. However, the ALK used in the reference model was considered to provide a reasonable proxy for growth trends, and the assessment results were robust to related uncertainties, as stock estimates were similar for all options explored (Figure 4.10).

Exploring the period of ALK data and growth deviations

The exploratory runs with different end years for ALK data and estimated growth deviations suggested that ALK data should be included for the years deviations are estimated. The run estimating deviations until 2018, but including ALK data only until 2012 had poorer convergence and produced peculiar recruitment estimates is some years (Figure 4.10). The runs including ALKs up to the present, but estimating deviations either up to 2018 or up to 2012 produced very similar results (Figure 4.9, 4.10).

It was concluded to maintain the option used in the reference run, i.e. ALK3 (based on BITS age readings) included for 1991–2018, and deviations estimated for all years, as none of the explored alternative ALK options improved the model.

4.6.3.4 Other exploratory runs with different configurations

A number of alternative configurations of the model were explored in addition to the ones with different input data described in the sections above. The runs presented in this section are described in the table below, together with the likelihood values and stock estimates for the final year:

Model	Landings	DevM	DevG	Maturity	Historical growth	Selectivity
Reference	1946-2018	2000	1991	Time variant	1946-1990	Logistic
BITS	1946-2018	2000	1991	Time variant	1946-1990	Logistic
dev2012	1946-2018	2000	2012	Time variant	1946-1990	Logistic
dirichlet	1946-2018	2000	1991	Time variant	1946-1990	Logistic
doubleN	1946-2018	2000	1991	Time variant	1946-1990	Logistic, except passive gear (double normal)
fixedMat	1946-2018	2000	1991	Fixed	1946-1990	Logistic
Ianelli	1946-2018	2000	1991	Time variant	1946-1990	Logistic
h082	1946-2018	2000	1991	Time variant	1946-1990	Logistic
h090	1946-2018	2000	1991	Time variant	1946-1990	Logistic
HistGrwTag	1946-2018	2000	1991	Time variant	1946-1950; 1951-1970	Logistic
HISTLAN	1925-2018	2000	1991	Time variant	1946-1990	Logistic
M&Growthdevest	1946-2018	2000	1991	Time variant	1946-1990	Logistic
recset	1946-2018	2000	1991	Time variant	1946-1990	Logistic
Roffset	1946-2018	2000	1991	Time variant	1946-1990	Logistic
ShiftR0	1946-2018	2000	1991	Time variant	1946-1990	Logistic
tagging	1946-2018	2000	1991	Time variant	1946-1990	Logistic

Model	Weighting	Regime shift	Steep	Spawning, settlement month	Q (fleets)
Reference	Francis	None	0.99	5,8	Estimated, except(BITSQ1&2, Larvae)
BITS	Francis	None	0.99	5,8	Estimated, except(Larvae)
dev2012	Francis	None	0.99	5,8	Estimated, except(BITSQ1&2, Larvae)
dirichlet	Dirichlet	None	0.99	5,8	Estimated, except(BITSQ1&2, Larvae)
doubleN	Francis	None	0.99	5,8	Estimated, except(BITSQ1&2, Larvae)
fixedMat	Francis	None	0.99	5,8	Estimated, except(BITSQ1&2, Larvae)
Ianelli	Ianelli	None	0.99	5,8	Estimated, except(BITSQ1&2, Larvae)
h082	Francis	None	0.82	5,8	Estimated, except(BITSQ1&2, Larvae)
h090	Francis	None	0.90	5,8	Estimated, except(BITSQ1&2, Larvae)
HistGrwTag	Francis	None	0.99	5,8	Estimated, except(BITSQ1&2, Larvae)
HISTLAN	Francis	None	0.99	5,8	Estimated, except(BITSQ1&2, Larvae)
M&Growthdevest	Francis	None	0.99	5,8	Estimated, except(BITSQ1&2, Larvae)
recset	Francis	None	0.99	Time variant	Estimated, except(BITSQ1&2, Larvae)
Roffset	Francis	1988	0.99	5,8	Estimated, except(BITSQ1&2, Larvae)
ShiftR0	Francis	1988	0.99	5,8	Estimated, except(BITSQ1&2, Larvae)
tagging	Francis	None	0.99	5,8	Estimated, except(BITSQ1&2, Larvae)

Model	TOTAL	Catch	Equil_catch	Survey	Length_comp	Age_comp	Parm_softbounds	Parm_devs	Tag_negbin	Tag_comp	Recruitment
Referencelihog	2451	1	252	116	685	1420	0.003	-193	0.0	0.0	169
BITSlihog	2331	1	215	93	684	1398	0.003	-204	0.0	0.0	142
Dev2012lihog	3010	0	240	121	872	1701	0.003	-85	0.0	0.0	160
Dirichletlihog	14116	0	326	120	2251	11374	0.003	-158	0.0	0.0	202
DoubleNlihog	2490	1	254	116	714	1427	0.010	-193	0.0	0.0	172
FixedMATlihog	2456	1	261	106	687	1416	0.003	-192	0.0	0.0	175
Ianellilihog	2713	1	242	75	1453	955	0.003	-181	0.0	0.0	167
Step082lihog	2689	0	293	114	693	1433	0.003	-192	0.0	0.0	349
Step090lihog	2585	0	282	114	689	1424	0.003	-193	0.0	0.0	267
HistGrwTaglihog	2458	1	229	118	690	1459	0.003	-191	0.0	0.0	151
HISTLANlihog	2019	0	0	85	669	1413	0.003	-186	0.0	0.0	38
MGrowth_devestlihog	2223	1	265	128	703	1438	0.003	-493	0.0	0.0	181
Recsetlihog	2454	1	254	116	686	1417	0.003	-193	0.0	0.0	172
Roffsetlihog	2067	1	28	44	669	1389	0.003	-178	0.0	0.0	114
ShiftR0lihog	2108	1	55	62	672	1397	0.003	-186	0.0	0.0	107
Tagslihog	4568	0	266	104	681	1438	0.003	-193	30	2064	177

Model	InitEQ_Regime	Forecast_Recruitment	Parm_priors	Crash_Pen	Gradient	R0	SSB2019	F2018	B0	SSB0	Recr0
Referencelihog	1.32E-32	2E-02	0	0	3E-02	14.7	81254	0.29	174192	831211	2530840
BITSlihog	0	4E-01	0	0	1E-02	14.6	49043	0.49	249611	798372	2210870
Dev2012lihog	0	3E-02	0	0	2E-05	14.7	83610	0.34	188508	867945	2457880
Dirichletlihog	0	2E-04	0	0	3E+02	14.7	91611	0.31	162923	886410	2430310
DoubleNlihog	0	2E-02	0	0	5E-03	14.7	82698	0.28	175916	833698	2304410
FixedMATlihog	0	3E-02	0	0	1E-02	14.7	64771	0.27	171735	889584	2323910
Ianellilihog	0	1E+00	0	0	1E+00	14.7	80396	0.28	192343	873184	2416250
Step082lihog	0	4E-05	0	0	5E+00	14.7	80573	0.28	185491	861849	2384090
Step090lihog	0	1E-03	0	0	1E-01	14.6	83400	0.29	171814	821967	2289700
HistGrwTaglihog	0	2E-02	0	0	9E-03	15.0	105482	0.25	196892	1145750	3139150
HISTLANlihog	0	8E-02	0	0	2E-01	14.6	70920	0.32	160372	802700	2196130
MGrowth_devestlihog	0	1E-02	0	0	3E-01	14.6	76226	0.33	149008	833221	2289280
Recsetlihog	0	2E-02	0	0	5E-01	14.7	81085	0.29	173199	833673	2304530
Roffsetlihog	3.59E-33	1E+00	0	0	1E-03	14.2	120592	0.23	87448	545847	1494260
ShiftR0lihog	0	6E-01	0	0	7E-03	14.3	99276	0.26	556787	593037	1625210
Tagslihog	0	7E-02	0.051	0	7E-03	14.7	115095	0.29	174723	948194	2323290

The results of these models and associated fits to BITS data are shown in Figure 4.11. The complete details of all models are available partially in the ICES Sharepoint and upon requests to the authors.

Some of the models explored (Step090, HistGrwTag and M&Growth_devest) could be considered as plausible alternative configurations of the reference model. However, they did not substantially improve the fitting of the model and the results were similar to the reference run.

The runs that introduced a regime shift in productivity (Roffset, ShiftR0) in 1988 when an overall regime shift in the eastern Baltic Sea has been documented, resulted in increased productivity estimates in later period. This was contrary to expectations based on biological knowledge, and these runs were therefore discarded.

The runs using the Dirichlet or Ianelli methods for data weighting were discarded as these did not converge, similar to the run with steepness at 0.82.

Extending the model further back in time with historical catches back to 1925 (HISTLAN) was attempted. This resulted in very high estimates of stock size in the early part of the time-series, which was not in line with the available knowledge.

The model including historical tagging data directly in Stock Synthesis model was only included as preliminary, as more work would be needed, e.g. on tag reporting rates and tag mortality.

The run with constant maturity through time was only included to see the effect of it, as the change in size at maturation is well documented.

Exploring BITS configuration

One of the issues that was extensively explored was the CV of BITS survey and consequent fitting of the model to BITS data. When estimating the CV of BITS surveys in the model (SS run- BITS), this resulted in very high CV being estimated and poor fit to the indices (Figure 4.11). In contrast, when giving the CV of BITS a low value (0.1), the model fits well to BITS, however compromising fit to some other components, especially the length compositions.

During the benchmark meeting, a number of exploratory runs were made, mainly aiming at investigating whether the fit to length composition data could be improved. It was pointed out that some discrepancy was present in the fit to length composition data with difficulties of the model to fit the highest peaks in commercial length composition and at the same time estimating slightly larger proportion of larger fish than observed in the data (Figure 4.22). Also, concerns were expressed by the group whether modelling of BITS indices entirely accounts for the changes in survey design around year 2000–2001.

The following exploratory runs were conducted:

Model	BITS Index	BITS length Comp	BITS catchability	BITS sample size
Reference	Full time-series	Full time-series	constant	number of hauls
BITS2000	2002 onwards	2002 onwards	constant	number of hauls
BITS2000Index	Full time-series	2002 onwards	constant	number of hauls
BITSoffset	Full time-series	Full time-series	separate from 2002	number of hauls
LowLsample	Full time-series	Full time-series	constant	reduced Q1

The run assuming different catchability in BITS from 2002 onwards as well as the run leaving out the earlier part of the BITS surveys (prior to 2002) resulted in almost flat stock dynamics from late 1990s onwards (Figure 4.12), which was not in line with observations from both the BITS survey as well as the SSB estimates from egg production method (Figure 4.2). Leaving out just the length composition in the early years or down weighing these data did not influence the results. None of these exploratory runs could improve the fit to length composition data. Including the BITS index in units of biomass instead of abundance was explored as well, and had no impact on the results or model fitting.

Therefore, the time-series and configuration of BITS were maintained as in the reference run. BITS CV was kept at 0.15 as the average of the time-series, as a compromise between reasonably fitting to the survey dynamics and at the same time not assuming an extremely low CV. The trends from BITS surveys are considered to reflect the true stock dynamics, being consistent between Q1 and Q4 surveys and also confirmed by independent estimates of SSB dynamics from the egg production method and by fisher observations.

Exploring selectivity

Exploratory runs were conducted assuming double-normal selectivity for all fleets, to possibly improve the fit to the length composition data as especially the passive gears often might have a dome shaped selectivity. The estimated selectivity was still following a logistic shape for all fleets, and the fits to length compositions were not improved by using double-normal selectivity. As logistic curve requires fewer parameters to be estimated, logistic selectivity was maintained, as in the reference run.

Exploring SigmaR

Exploratory runs were made with different σ_R values (which was set to 0.6 in the reference run). Alternative values for σ_R had minor impact on the assessment results, unless a very low value was used (Figure 4.13; see also likelihood profiles of σ Figure 4.25).

Exploring age-error

The reference model does assume a very low age imprecision and no bias in ageing. During the benchmark meeting, adding an age-imprecision matrix in the Stock Synthesis model was explored. Age reading of Eastern Baltic cod is associated with uncertainties, thus the use of an age-error matrix would be appropriate. Different otolith exchange exercises have been conducted in the past, giving an estimate of variability of age readings. Identifying the bias in age readings is however more difficult, as there are currently no known-age samples available for Eastern Baltic cod. A preliminary run was made during benchmark including imprecision in the age information. Two different levels of imprecision were applied, one for the historical and one for the more recent period. The results showed somewhat improved fit to the length composition data, especially for Active gears, and estimated both higher growth and higher M, with some effects on stock estimates, though not changing the overall perception of the stock status. As the values included in age imprecision matrix during benchmark were considered preliminary and had not been evaluated at the data meeting, the run with ageing imprecision was only considered as exploratory. It was recommended that future work should be directed towards working up an age-imprecision and bias matrix by year (section 4.11).

4.6.4 Final Stock Synthesis run and diagnostics

Overview of the datasets included in the final Stock Synthesis model is shown in Figure 4.14 and Table 4.1. The settings and estimated parameters are presented are Table 4.2. The estimated biological changes (time variant natural mortality and growth), and estimated time invariant selectivity are shown in Figures 4.15-4.18. The estimated deviations in M and growth are in line with the available biological information (e.g. ICES WKBEBCA, WKIDEBCA).

Residuals for length compositions show a pattern of underestimating the peak in length distribution and slightly overestimating the proportion of the larger cod (Figures 4.19 and 4.21), however the residuals are generally small. For most fisheries, there is a reasonable overall fit to the length and age composition data. Annual fits to length and age compositions and to ALKs are provided in WD7.

Overall, the model provides a reasonable fit to the trends in the CPUE indices (Figure 4.22). The model fit to the BITS surveys indices was good except for the 2008–2011, which were always underestimated in the model. A non-random pattern of residuals may indicate that some heteroscedasticity is present, or there is some leftover serial correlation (serial correlation in sampling/observation error or model misspecification). Several well-known nonparametric tests for randomness in a time-series include: the runs test, the sign test, the runs up and down test, the Mann-Kendall test, and Bartel's rank test (Gibbons and Chakraborti, 1992). Here we used the runs test to evaluate whether residuals of the surveys, in particular the BITS surveys are random over time, because this test has been used to diagnose fits to indices and other data components in assessment models (e.g. SEDAR 40, 2015). The results of the runs test are presented in Figure 4.23. The runs test indicated that the fit of the CPUE indices is adequate and very few residuals larger than 1, indicating a random pattern of the surveys residuals.

Comprehensive diagnostics of the model is described below:

Retrospective analyses

Retrospective analysis is a diagnostic approach to evaluate the reliability of parameter and reference point estimates and to reveal systematic bias in the model estimation. It involves fitting a stock assessment model to the full dataset. The same model is then fitted to truncated datasets where the data for the most recent years are sequentially removed. The retrospective analysis was conducted to the reference model for the last 5 years of the assessment time horizon to evaluate whether there were any strong changes in model results.

Given that the variability of Mohn's index depends on life history, and that the statistic appears insensitive to F, Hurtado-Ferro (2014) proposed the following rule of thumb when determining whether a retrospective pattern should be addressed explicitly. Values of Mohn's index higher than 0.20 or lower than -0.15 for longer-lived species (upper and lower bounds of the 90% simulation intervals for the flatfish base case), or higher than 0.30 or lower than -0.22 for shorter-lived species (upper and lower bounds of the 90% simulation intervals for the sardine base case) should be cause for concern and taken as indicators of retrospective patterns. However, Mohn's index values smaller than those proposed should not be taken as confirmation that a given assessment does not present a retrospective pattern, and the choice of 90% means that a "false positive" will arise 10% of the time. In both cases, model misspecification would be correctly detected more than half the time. The retrospectives of the reference model were rather stable (Figure 4.24). The estimated Hurtado-Ferro (2014) variant of the Mohn´s index was small for SSB (0.08) and F (0.13) but larger than the threshold for R (0.60), which is expected as it takes about 2-3 years of data for a year class to be determined with high precision as shown by the squid plot of retrospectives of recruitment deviations (Figure 4.24).

Likelihood profiles

Likelihood profiling is an automated routine in Stock Synthesis, that allows to evaluate model performance across a range of values of an input parameter (generally R0, σ_R and steepness). Here we performed the likelihood profile of R_0 and σ_R for the reference model (Figure 4.25). The likelihood profile of R0 shows a minimum at the model estimated minimum, although there is an apparent conflict between the different components in the estimate of R_0, especially age and length compositions compared to survey indices.

The likelihood profile of σ_R shows a minimum at the model estimated minimum (0.6) with no apparent conflict between the different data components. σ_R is the stochastic recruitment process error and the estimation of this parameter within integrated models is generally recognized to be problematic (Kolody *et al.*, 2019) so that σ_R individual recruitment estimates analogous to traditional VPA. A meta-analysis of the estimation of σ_R done outside the operative model (ISSF 2011) yielded a median estimate between 0.2 and 0.5, which suggested that σ_R is often inflated in assessment models. However, we also investigated the effect of different assumed values for σ_R and we found that using smaller or larger values of σ_R has a little effect on the estimates stock trends and absolute values (Figure 4.13), which is reassuring that the management advice is not affected by the choice of the σ_R value.

Hindcasting

A major uncertainty in stock assessment models is the difference between model estimates and reality. The validation of model prediction is difficult, however, as fish stocks can rarely be observed and counted. Kell *et al.* (2016) showed how hindcasting can be used to evaluate model prediction skill. When conducting hindcasting, a model is fitted to the first part of a time-series and then projected over the period omitted in the original fit. Prediction skill can then be evaluated by comparing the predictions from the projection with the observations using for example the mean absolute scaled error (MASE; Hyndman, R. J. and Athanasopoulos, G. (2013)) indicator. Hindcasting was conducted for the reference model (Figure 4.26). The results showed that a substantial hindcasting error is present for most of the surveys, with MASE values generally much larger than the 1.0 threshold. However, hindcasting was much less severe for the SSB within a 3 years' span, which is the time frame used for the forecast. This indicated that, although the model is not able to predict the survey indices in future, the prediction of the SSB, which is the main measure of interest for the management is rather good. However, it is important to highlight that the hindcasting of SSB is good only within the 3 years span but that the model deteriorates after 4 to 5 years. This means that this particular models configuration is appropriate to providing short-term advice but that it is desirable that the model is benchmarked again in 4 to maximum 5 years.

Jittering

The jitter procedure helps to verify the stability of the model examining the effect of varying the starting values of the model input estimated parameters on model results. An accurate model should converge on a global solution across a reasonable range of starting values input parameters.

In this case, 100 runs were performed considering a 10% of jitter of the initial parameters, which means that a small random jitter is added to the initial parameter values. Starting values are jittered based on a normal distribution based on the pr(PMIN) = 0.1% and the pr(PMAX) = 99.9%. Results (Figure 4.27) confirmed the stability of the model and the absence of local minima. It is however important to stress that the absence of a local minima when running jittering is not a guarantee that the model is not indeed stuck in a local minimum, although its absence reduced the risks that this occurs (Subbey, 2018).

MCMC

Markov chain Monte Carlo (MCMC) methods comprise a class of algorithms for sampling from a probability distribution. By constructing a Markov chain that has the desired distribution as its equilibrium distribution, one can obtain a sample of the desired distribution by observing the chain after a number of steps. The more steps there are, the more closely the distribution of the sample matches the actual desired distribution. Markov chain Monte Carlo methods create samples from a possibly multidimensional continuous random variable, with probability density proportional to a known function. These samples can be used to evaluate an integral over that variable, as its expected value or variance. Practically, an ensemble of chains is generally developed, starting from a set of points arbitrarily chosen and sufficiently distant from each other. Those are then used to estimate the posterior distribution of the parameters of interest within the model. For Eastern Baltic cod, we run an MCMC with 100 000 iterations, with no burn-in period and thinning each 1000 iterations. The results showed that the MCMC is rather similar to the MLE estimated, which is another indication of the robustness of the model (Figures 4.27 and 4.28). This is also confirmed by the Geweke and Heidelberger and Welch statistic (Figure 4.28).

4.7 Exploratory stock assessment: SPICT

SPICT (Stochastic Production model in Continuous Time) has previously been used for Eastern Baltic cod to classify the stock status in relation to proxy MSY reference points. SPICT is a standard method used for data-limited stocks, it is extensively tested, reviewed and endorsed by ICES working groups (WKLIFE, WKPROXY,WKMSYCat34).

SPICT estimates for Eastern Baltic cod have only been used as relative (B/B_{MSY}, F/F_{MSY}) in former years. The aim of this benchmark was to establish an assessment with absolute stock estimates (ICES category 1). As Stock Synthesis was accepted by the group to be used as the basis for advice for Eastern Baltic cod, SPICT was not thoroughly evaluated at this benchmark. However, it was recommended to maintain SPICT in WGBFAS as a secondary model for Eastern Baltic cod.

New developments in SPICT model for Eastern Baltic cod were presented at benchmark. These included allowing for a gradual change in productivity over time, instead of different production curves (regimes) at fixed time points (Figures 4.30 and 4.31). The change in productivity estimated from SPICT is in line with the results from Stock Synthesis. Additionally, Thorson's prior was included as this prior information gives a more robust assessment (e.g. less retrospective patterns). Thorson's prior increases stability and reduces uncertainty in the model. Overall, the assessment with SPICT performs reasonably well – with relatively low uncertainty and no retrospective pattern. SPICT will be continued to be explored in WGBFAS as a secondary model for the Eastern Baltic cod.

4.8 Appropriate Reference Points (MSY)

Biomass reference points

B_{lim} is a biomass limit below which a stock is considered to have reduced reproductive capacity. ICES generally defines B_{lim} by looking at the stock–recruitment relationship and identifying the spawning-stock biomass (SSB) level below which recruitment reduces with SSB, e.g. the change point of a segmented regression.

For Eastern Baltic cod, it is well recognized that large changes have occurred both in environmental conditions and in the distribution of cod and therefore restricted period (from late 1980s) has been used for the stock recruitment (S-R) relationship when defining reference points in the past (e.g. ICES WKBALT 2013). WKBALTCOD2 recognized that major ecological changes have additionally occurred in later years, which need to be taken into account when setting B_{lim},. This implies that it is no longer relevant to consider the entire time-series from the late 1980s onwards for S-R, as has been done in the past, and the WKBALTCOD2 therefore focused on evaluating the reproductive capacity of the stock in most recent years.

This is because of the following reasons:

The SSB in later years is not only reflecting the dynamics in stock size, but is additionally strongly influenced by the reduced size at maturation (Figure 4.32). The SSB in recent years contains a large proportion of small individuals that were not yet part of SSB in former years (before 2000s). The biomass of the relatively larger cod that formed the spawning stock before the 2000s is currently at a historic low level (Figure 4.32). The eggs of young female Eastern Baltic cod have considerably lower survival at poor hydrographic condition compared to the eggs of older females (Hinrichsen *et al.*, 2016). Furthermore, the condition of spawners has much deteriorated in later years, due to low nutritional condition and high infestation with parasites. Thus, the reproductive capacity of a specified amount (tons) of SSB today (consisting of small individuals at poor condition) is likely not equal to the reproductive capacity of the same amount of SSB in the past.

The biological characteristics of Eastern Baltic cod likely to influence its reproductive capacity have gradually deteriorated since the 1990s, with some levelling off in the latest years (Figure 4.33). Therefore, the WKBALTCOD2 focused on evaluating the reproductive capacity of the stock for the most recent years. The year classes from 2015 and 2016 are estimated to be among the lowest since the 1990s (Figure 4.34). Preliminary information from the BITS Q4 2018 survey indicates a weak year class also for 2017. Moreover, preliminary information from the 2018 ichthyoplankton surveys shows very low larval abundances throughout the spawning season, suggesting a poor year class also for 2018. This sequence of poor year-classes raises concerns about the current reproductive capacity of the stock, the recruitment possibly being impaired by the quality of the spawning stock. Therefore, the size of spawning stock (SSB) in tons alone is not considered representative for reproductive capacity for the stock at present, as the quality of the SSB needs to be considered as well.

It was concluded that B_{lim} should currently not be set lower than the most recent SSB that was still able to produce a strong year class, when much of the adverse developments affecting the quality of the SSB had already taken place. The latest relatively strong year class was formed in 2012 from an SSB of 98 000 t (Figure 4.34). Therefore, B_{lim} was set to this level, i.e. 98 000 t.

Due to the presently very dynamic biological situation for the Eastern Baltic cod, the current B_{lim} at 98 000 t is considered to be applicable only in short term. The reproductive capacity of the stock needs to be closely monitored in coming years, and when new information becomes available, the B_{lim} value needs to be re-evaluated.

B_{lim} at 98 000 t corresponds to B_{pa} at 124 000 t ($B_{lim} \times \exp(1.645 \times \sigma)$, where σ=0.14).

Estimation of F_{MSY}

The Eastern Baltic cod stock experiences large changes in productivity, which questions the applicability of the F_{MSY} concept for this stock that assumes long-term equilibrium.

The estimation of F_{MSY} was attempted using the ICES standard software Eqsim. The biology (weights, natural mortality, maturity) and selectivity were based on the latest years (2015-2018). For stock–recruitment, the hockey-stick function was applied, with a breakpoint at B_{lim}. The years before 1986 were excluded in the SR time-series, in line with previous assessments (e.g. ICES WKBALT 2013), due to large changes in environmental conditions and in the distribution of cod, with only one functioning spawning ground since the mid-1980s and the stock being concentrated in the southern Baltic in later years.

It is recognized that due to the large biological changes (poor condition, small size at maturation, few larger individuals on the stock) in later years, only most recent years should be considered for S-R (see the section above on biomass reference points). For technical reasons, the time-series of S-R from 1986 onwards was used in this exercise to estimate F_{MSY}, which is however not expected to invalidate the Eqsim results, as the average level of R at SSB >B_{lim} was in line with the average recruitment in the latest years (after 2012) (Figure 4.34).

The Eqsim analyses showed that even with F_{MSY} at 0 the SSB would not be kept above B_{lim} (98 000 t) in the long term, with 95% probability.

The sensitivity of this result to the B_{lim} value was also explored. Even when applying a substantially lower value for B_{lim} (53 000 t), the result in terms of the stock being below B_{lim} with more than 5% probability even at F_{MSY} = 0 remained unchanged. In other words, following the ICES MSY framework for this stock, the estimated F_{MSY} is equal to 0. For this reason, no F reference points were defined for this stock.

4.9 Short-term projections

The short-term projections were done with Stock Synthesis. The assessment period is set to the last year of the survey, which is one year later than then last year for which catches are available. Therefore, to be able to use the latest survey information in the assessment, catches for that year need to be assumed (corresponds to the intermediate year in the forecast). It is to be decided at WGBFAS on an annual basis, what is the reasonable assumption for catches in this intermediate year. Due to large changes in biology currently ongoing, it is also to be decided at WGBFAS whether biology (natural mortality, growth) in the forecast should be based on the latest year only, or average of the last few years. Recruitment in the forecast period was decided to be set to the average from 2013 until the last year in the assessment time-series for which recruitment deviations are estimated in the Stock Synthesis model. This currently corresponds to the low-average recruitment, i.e. not including in recruitment predictions the latest relatively strong year classes from 2011–2012.

As there is no F reference point for this stock, probabilistic forecast with MCMC was proposed to be used instead. In this approach, catch and SSB levels corresponding to different F factors are calculated as in typical deterministic short-term forecast but using MCMC to make it possible to also include the associated probability/risk of the SSB to be below B_{lim} and $B_{trigger}$ for each year of forecast. At the benchmark, this approach was approved to be used, and the actual forecast will be performed in the next WGBFAS.

4.10 Final remarks on quality of the established analytical assessment

The assessment of the Eastern Baltic cod has since 2014 been based on ICES data limited approach. The main reason for rejecting the last analytical stock assessment was the strong retrospective bias observed in the assessment. Possible explanations for this bias included issues with age reading and possible changes in growth and/or natural mortality that were not accounted for in the model. Also, it was unclear whether and to what extent such changes had taken place. In the following years, research on these matters was intensified, and regular meetings were held at ICES (WKSIBCA, WKBEBCA, WKIDEBCA) to collate and synthesize the new scientific knowledge, and discuss its use for stock assessment purposes.

In contrast to the situation in 2014 when the former analytical stock assessment was rejected, there is now ample evidence of the reduced growth of the Eastern Baltic cod from different investigations (summarized in WKBEBCA, WKIDEBCA). In addition, natural mortality of the stock is expected to have increased (WKBEBCA, WKIDEBCA). In the established stock assessment using Stock Synthesis, the magnitude of change in growth from 1991 onwards is estimated within the model using yearly age–length keys, which are based on traditional age readings. It is recognized that age readings for the Eastern Baltic cod are uncertain, especially for later years, while age imprecision is not explicitly accounted for in the model. However, the advantage of the approach used in Stock Synthesis compared to previous age based assessment is that the changes in the von Bertalanffy growth parameters are estimated internally within the model. In former age based assessments, each country provided their catch-at-age numbers externally. Thus, the assessment was essentially operating with a number of different national growth curves. However this was not explicit and it as not possible to retrieve or analyse the underlying growth parameters, because only the final product in terms of national catch-at-age numbers was available.

In the present approach used in Stock Synthesis, different labs' age readings are contributing to a joint ALK that is used to inform the trend in common growth parameters for the stock within the model. Thus, the ALK is not applied directly and solely on catch at length, but it is used to estimate the yearly values and thus the trend in von Bertalanffy growth parameters, which are thereafter used to derive catch-at-age from catch at length information. Using the ALKs in any-way is of course still problematic, if the age information is unreliable. Therefore, the following steps were considered to ensure that the ALKs currently used in Stock Synthesis provide a reasonable proxy for estimating changes in growth:

i. The approach used in Stock Synthesis, where changes in von Bertalanffy growth parameters are estimated internally in the model allows validating the signal in the ALK data, and subsequently estimated change in growth parameters with additional biological knowledge.
ii. The estimated change in growth is in line with observed changes in biology of the stock and environmental conditions (see section 4.6.3.3).
iii. The estimated change in growth is in line with preliminary growth information from recent tagging program (see section 4.6.3.3).
iv. The exact values for von Bertalanffy growth parameters are associated with uncertainties due to imprecise age information. This is affecting also natural mortality estimates, as growth and M are confounded. However, the results of stock assessment in terms of stock status were found to be robust to these uncertainties associated with separating between M and growth (Figure 4.10).

For these reasons, the ALKs currently used are considered to provide a reasonable proxy for informing growth for stock assessment purposes. This is considered a temporary solution, as an alternative method for estimating growth is being developed (see section 4.11).

The established assessment with Stock Synthesis shows no retrospective bias, fits reasonably to the data, performs well in different other diagnostics (see section 4.6.4) and appears to be robust to the uncertainties related to separating between M and growth. Therefore, the quality of the assessment was found reasonable and the assessment appropriate to be used as the basis for advice for the Eastern Baltic cod.

4.11 Future research and data requirements

Future work should focus on improving the growth estimates, which will allow separating between growth and natural mortality with greater precision. Progress in this area is expected after the final results from the ongoing TABACOD project become available. This is expected to provide refined growth estimates from recent tagging, as well as a new method for growth determination based on otolith microchemistry, that can be used to monitor the growth of Eastern Baltic cod in future. The validated growth information that is expected to become available from this work can also contribute to developing an age-error matrix to account for past uncertainties in age information in Stock Synthesis model. Development of age error matrix is seen as one of the major focus area for future improvement of the present Stock Synthesis model for Eastern Baltic cod.

Additionally, Stock Synthesis model would benefit from information on sample size associated with length distributions of commercial catches. Currently no meaningful measure on sample size is available, and the data for all years, fleets and quarters are treated equally. This could be improved, if a meaningful measure representing sample size of combined international commercial data could be developed.

4.12 Figures and Tables

Note: In all figures showing spawning-stock biomass from Stock Synthesis, the SSB refers to females only.

Table 4.1 Eastern Baltic cod. Input data of the final Stock Synthesis model.

Type	Name	Year range	Range	Time variant
Catches	Catch in tonnes split into Active/Passive and quarters (see section 4.5.1)	1946–2018	0 - 15+	
Age compositions	Catch in numbers per age class of the fleets, by Q	1946–2006	0 - 12+	
Length compositions	Catch in numbers per length class of the fleets, by Q, and BITS Q1 and Q4	2000–2017(2018)	5 – 120 cm	
Maturity ogives	Size at 50%maturity(L50) and slope	1946–2018		Yes (1998-2018, Lmat)
Growth	von Bertalanffy growth curve	1946–2018		Yes (1991-2018)
Natural mortality	Natural mortality by age class	1946– 2018	0 - 15+	Yes (2000-2018)
Age length compositions	Age length keys from BITS Q1 and Q4	1991–2018	0 – 12+	Yes (1991-2018)
Surveys indices	CPUE from BITS Q1, Q4, and trawl surveys 1 and 2	1975–2018		
Commercial CPUE indices	Commercial CPUE 1-3	1948–1989		
SSB index	SSB from egg production method	1986–2017		
Larval index	Larval abundance	1987–2017		

Table 4.2. Eastern Baltic cod. Settings and results of the reference model. The table columns show: number of estimated parameters, the initial values (from which the numerical optimization is started), the intervals allowed for the parameters, the priors used, and the value estimated by maximum likelihood. Parameters in bold are set and not estimated by the model.

Parameter	Number estimated	Initial value	Bounds (low,high)	Prior	Value (MLE)
Natural mortality (age classes 0.5, 1.5, 5.5, 15.5)		1.243, 0.857, 0.361, 0.215			
M (2000-2018) of age class 5.5	19	Estimated using random walk annual deviations	(0.1,2.0)	no prior	0.361-0.82
Stock and recruitment					
Ln(R_0)	1	14.8	(13,16)	no prior	14.75
Steepness (h)		0.99			
Recruitment variability (σ_R)		0.60			
Ln (Recruitment deviation): 1946-2016	71				
Recruitment autocorrelation		0			
Growth					
L_{inf} (cm) (1946-1990)		125.27			
L_{inf} (cm) (1991-2018)	28	Estimated using random walk annual deviations	(40-150)	no prior	(125.27-57.2)
k (1946-1990)		0.10			
k (1991-2018)	28	Estimated using random walk annual deviations	(0.07-0.45)	no prior	(0.1-0.18)
L at minimum age (0.5 years) t_0		12			
CV of young individuals	1	0.290	(0.05-0.8)	no prior	0.27
CV of old individuals		0.05			
Weight (kg) at length (cm)					
a (1946-1990)		6.58e-06			
b (1946-1990)		3.1553			
a (1991-1993, 1994- 1996, 1997-1999, 2000 -2002, 2003-2005, 2006-2008, 2009-2011, 2012-2014, 2015-2018)		6.58E-06, 8.05E-06, 6.81E-06, 6.78E-06 6.76E-06, 7.47E-06 6.70E-06, 7.73E-06 8.90E-06			

Parameter	**Number estimated**	**Initial value**	**Bounds (low,high)**	**Prior**	**Value (MLE)**
b (1991-1993, 1994- 1996, 1997-1999, 2000 -2002, 2003-2005, 2006-2008, 2009-2011, 2012-2014, 2015-2018)		3.1213, 3.071, 3.033 3.0665, 3.0851, 3.0632 3.057, 3.0076, 3.0076			
Maturity					
Length (cm) at 50% mature (1946-1990)		38			
Slope of the length at maturity ogive		-0.23			
Length (cm) at 50% mature (1991-1997, 1998-2000, 2001-2007, 2008-2014, 2015-2018)		38, 36, 31, 26, 21			
Initial fishing mortality					
Active gears		0.60			
Selectivity (logistic)					
Active gears					
Time-invariant length based logistic selectivity	2	35; 12.68	(20,45; 0.01,50)	no prior	(39.1; 8.3)
Passive gears					
Time-invariant length based logistic selectivity	2	35; 10	(20,65; -12,15)	no prior	(42.2; 8.8)
BITS Q1 survey					
Time-invariant length based logistic selectivity	2	25,10	(15,50; -12,15)	no prior	(29.0;11.3)
BITS Q4 survey					
Time-invariant length based logistic selectivity	2	25,10	(15,50; -12,15)	no prior	(28.8; 10.6)
Commercial CPUE 1-3		Mirror active fleet			
Trawl surveys 1-2		Mirror BITS Q1			
Catchability					
BITSQ1					
Ln(Q) – catchability		Float option used			
Extra variability added to input standard deviation		0.01			
BITSQ4					

Parameter	**Number estimated**	**Initial value**	**Bounds (low,high)**	**Prior**	**Value (MLE)**
Ln(Q) – catchability		Float option used			
Extra variability added to input standard deviation		0.01			
Trawl survey 1					
Ln(Q) – catchability		Float option used			
Extra variability added to input standard deviation	1	0.1	(0.0,0.8)	no prior	0.297
Trawl survey 2					
Ln(Q) – catchability		Float option used			
Extra variability added to input standard deviation	1	0.1	(0.0,0.8)	no prior	0.015
Commercial CPUE 1					
Ln(Q) – catchability		Float option used			
Extra variability added to input standard deviation	1	0.1	(0.0,0.8)	no prior	0.10
Commercial CPUE 2					
Ln(Q) – catchability		Float option used			
Extra variability added to input standard deviation	1	0.1	(0.0,0.8)	no prior	0.07
Commercial CPUE 3					
Ln(Q) – catchability		Float option used			
Extra variability added to input standard deviation	1	0.1	(0.0,0.8)	no prior	0.31
Egg biomass index					
Ln(Q) – catchability		Float option used			
Extra variability added to input standard deviation	1	0.1	(0.0,1.2)	no prior	0.46
Larvae index					
Ln(Q) – catchability		Float option used			
Extra variability added to input standard deviation		0.3			

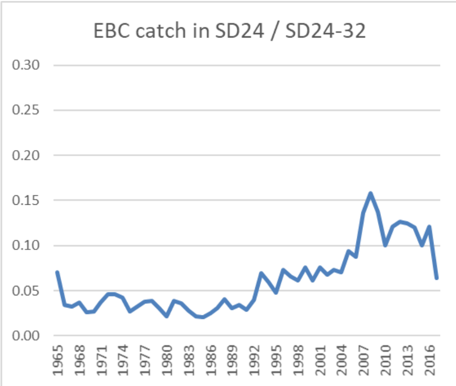


Figure 4.1 Proportion of the total Eastern Baltic cod catch (in SD 24-32) taken in SD 24.

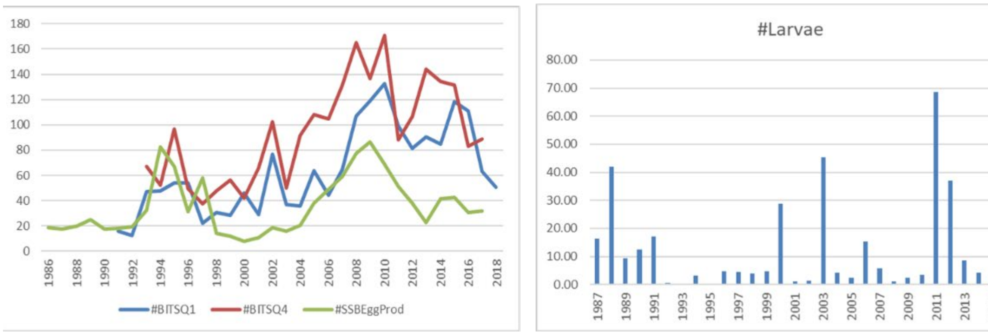


Figure 4.2. Time-series of tuning indices: BITS Q1 and Q4 (in abundance) and SSB index based on egg production method (in biomass), and index of the abundance of larvae.

Figure 4.3. Natural mortality-at-age applied for years 1946–1999.

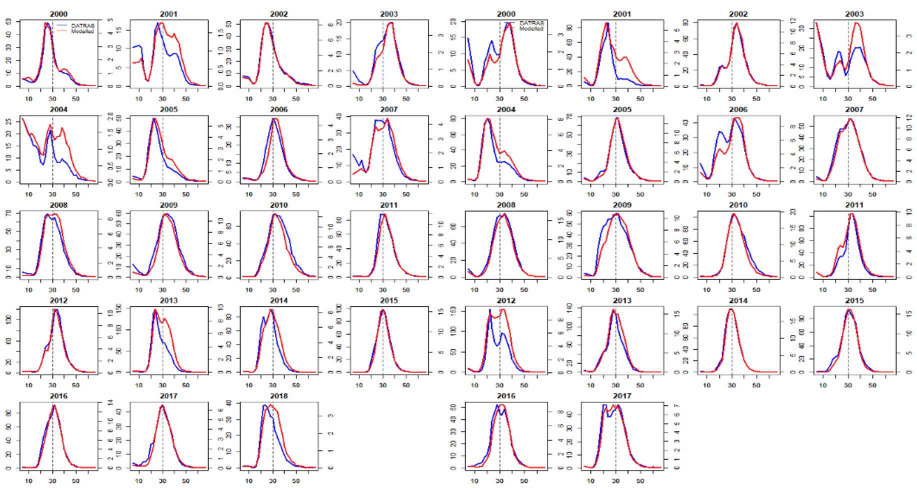


Figure 4.4. Length distributions of modelled BITS indices compared to those from DATRAS in Q1 (left panels) and Q4 (right panels)(data for SD25-32).

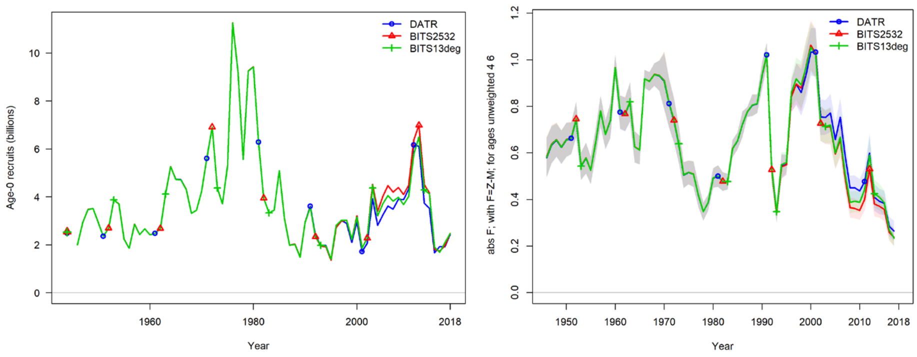


Figure 4.5. Recruitment and F estimates from Stock Synthesis runs with BITS indices from DATRAS (DATR) for the area SD 25-32 compared to modelled indices for SD 25-32 (BITS 2532) and modelled indices including a fraction from SD 24 (BITS13deg).

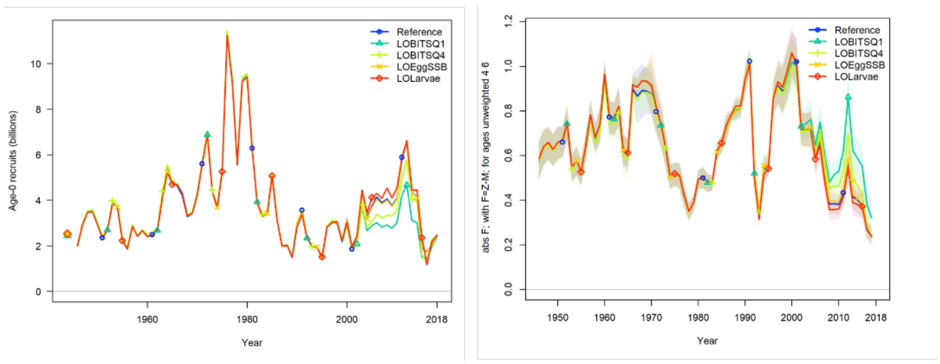


Figure 4.6. Recruitment and F estimates from Stock Synthesis runs leaving out one survey time-series at a time (BITS Q1 and Q4, SSB index based on egg abundance and Larval index).

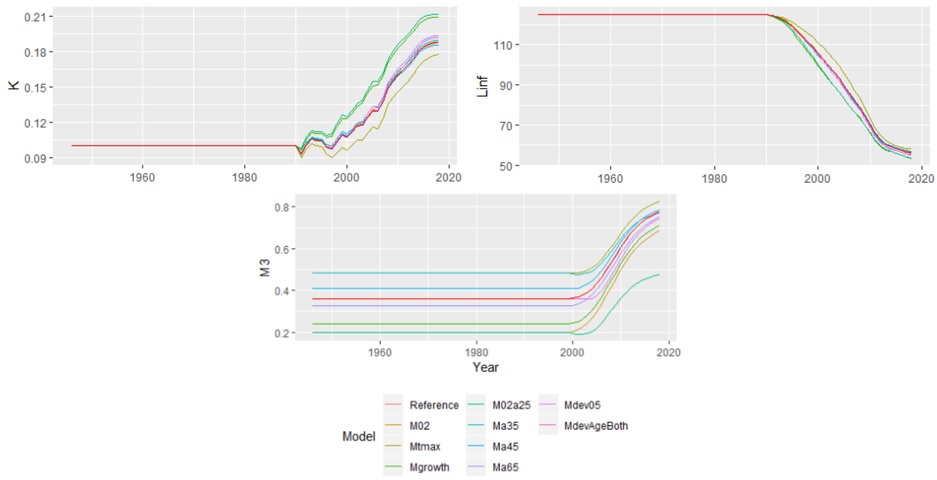


Figure 4.7. Estimated deviations in natural mortality, and growth parameters from exploratory runs with different M configurations (see section 4.6.3.2 for explanation of the runs).

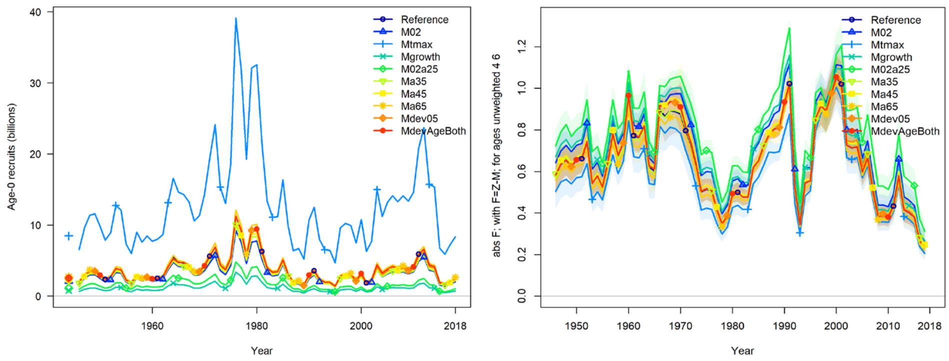


Figure 4.8. Recruitment and F estimates from Stock Synthesis runs with different M values and configurations (see section 4.6.3.2 for explanation of the runs).

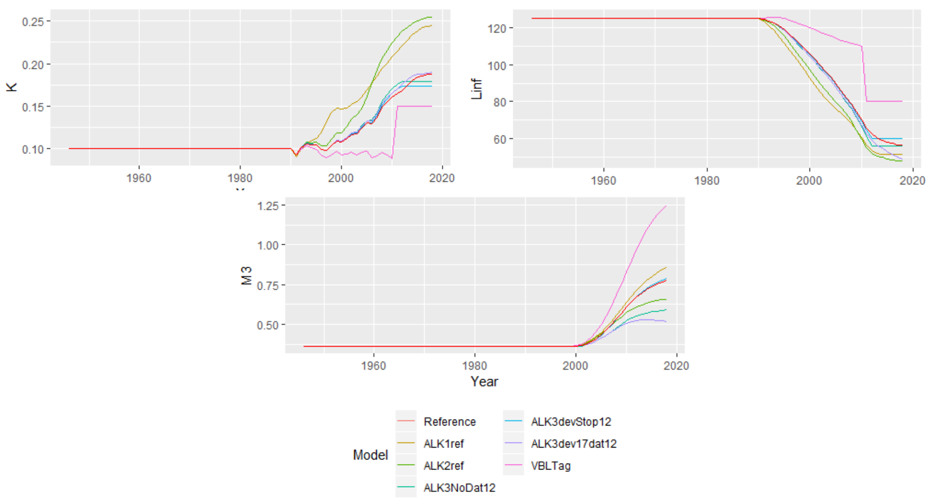


Figure 4.9. Estimated deviations in natural mortality and growth parameters from exploratory runs with different growth information (age–length key) (see section 4.6.3.3 for explanation of the runs).

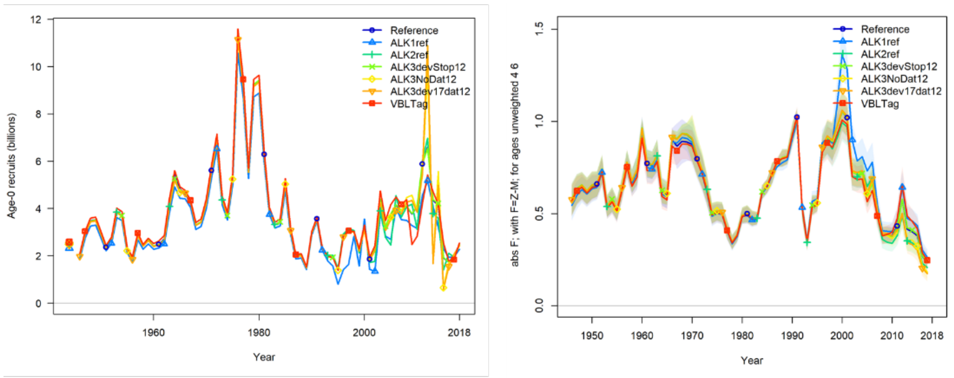


Figure 4.10. Recruitment and F estimates from exploratory runs with different growth information (age–length key) (see section 4.6.3.3 for explanation of the runs).

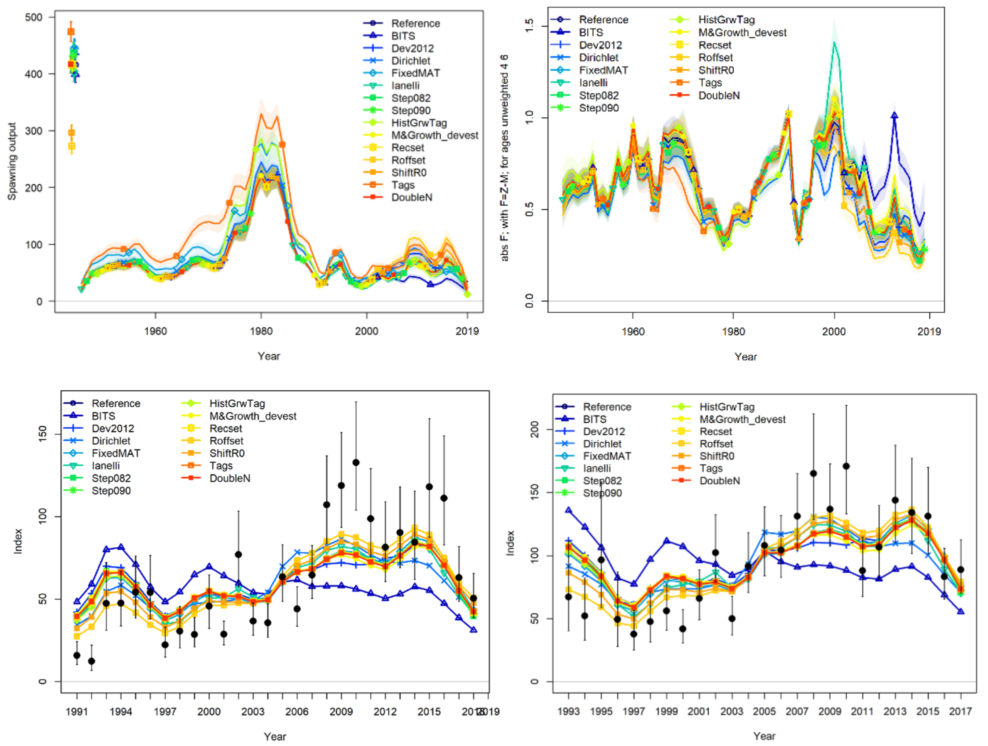


Figure 4.11. SSB (females only) and F estimates (upper panels) and fits to BITS (lower panels) from exploratory runs with different configurations (see section 4.6.3.4 for explanation of the runs).

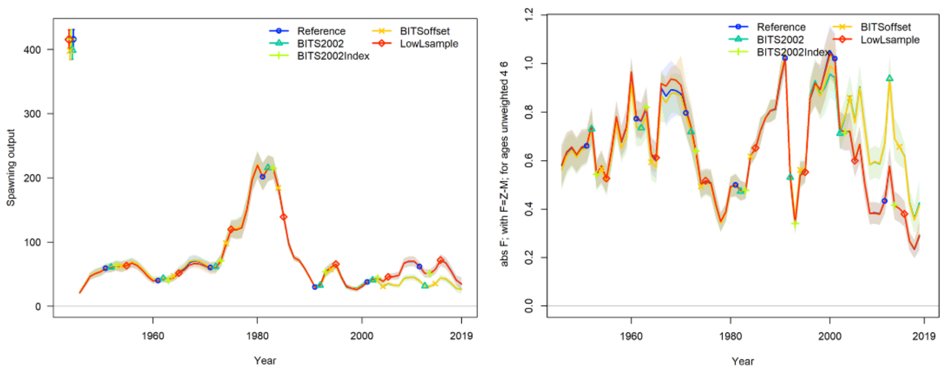


Figure 4.12. SSB (females only) and F from exploratory runs with different treatment of BITS data (see section 4.6.3.4 for explanation of the runs).

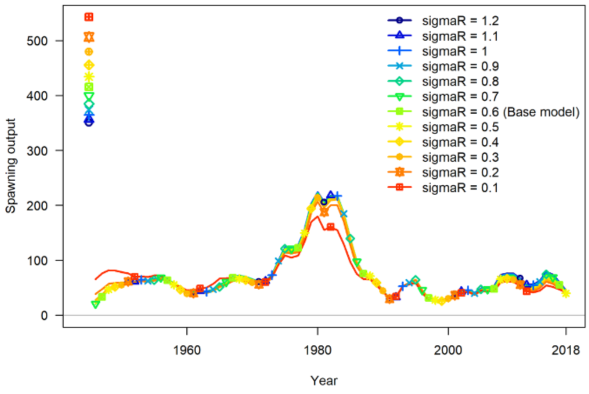


Figure 4.13. SSB (females only) estimates from exploratory runs with different sigmaR values.

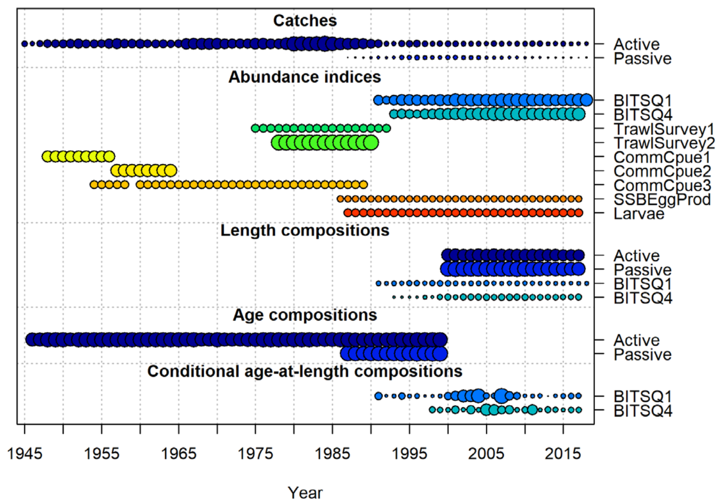


Figure 4.14. Summary of the input time-series included in Stock Synthesis.

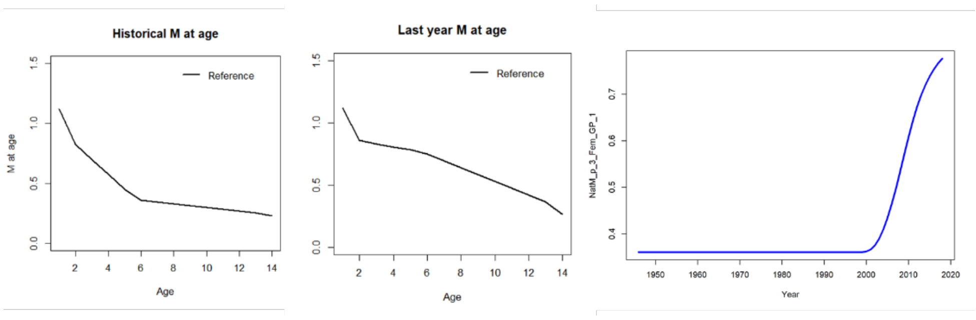


Fig 4.15. Left panel: M at age in historical period (1946-1999); middle panel: M at age estimated for last year in the assessment; right panel: M at age 5.5 with the estimated deviation from 2000 onwards.

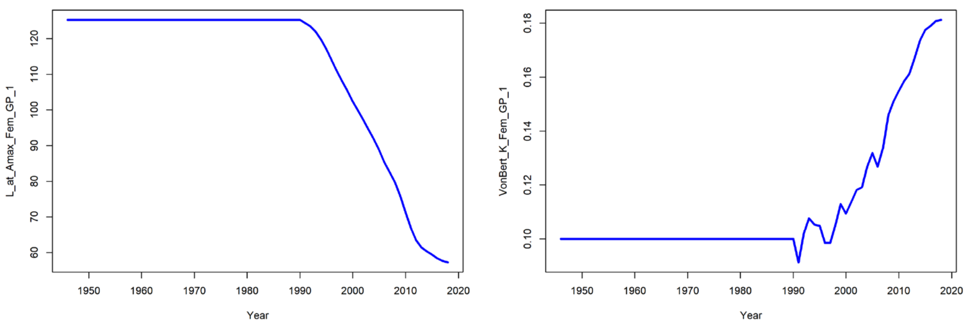


Figure 4.16. Estimated deviations in growth parameters (Linf-left panel and K- right panel) in the final Stock Synthesis run.

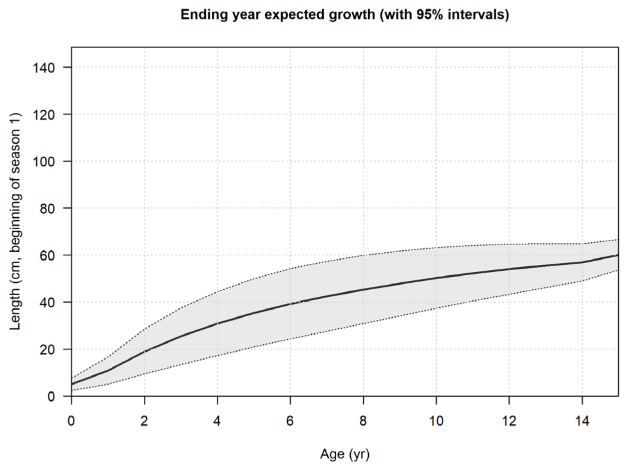


Figure 4.17. Estimated length-at-age in the final year of the assessment.

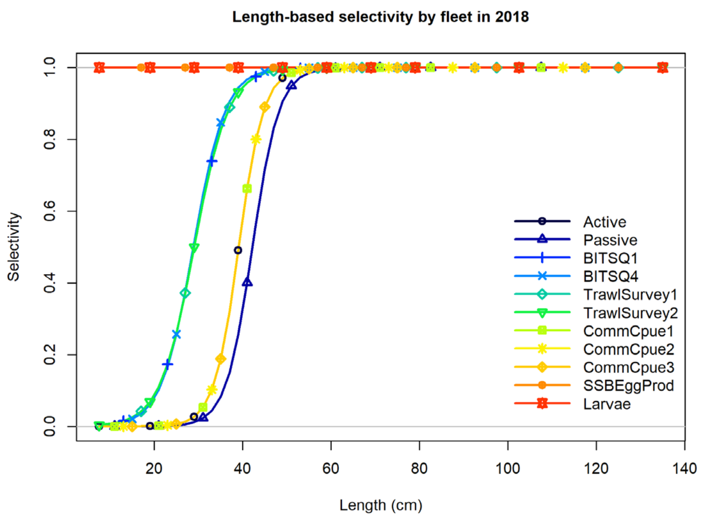


Figure 4.18. Selectivity estimates for different fleets.

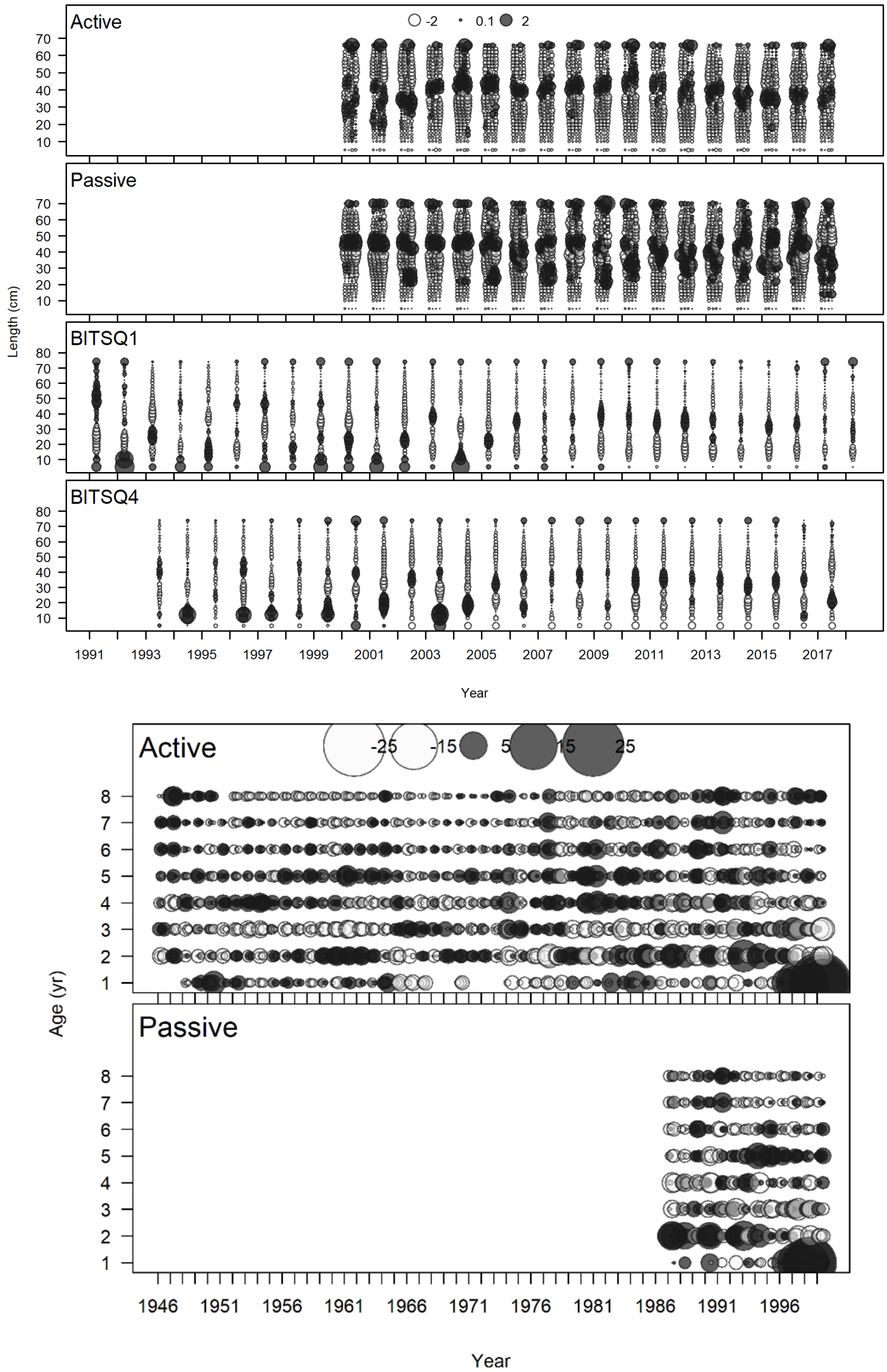


Figure 4.19. Residuals of fits to length (upper panels) and age (lower panels) composition data for different fleets.

Figure 4.20. Model fits to age composition data (combined across years and quarters).

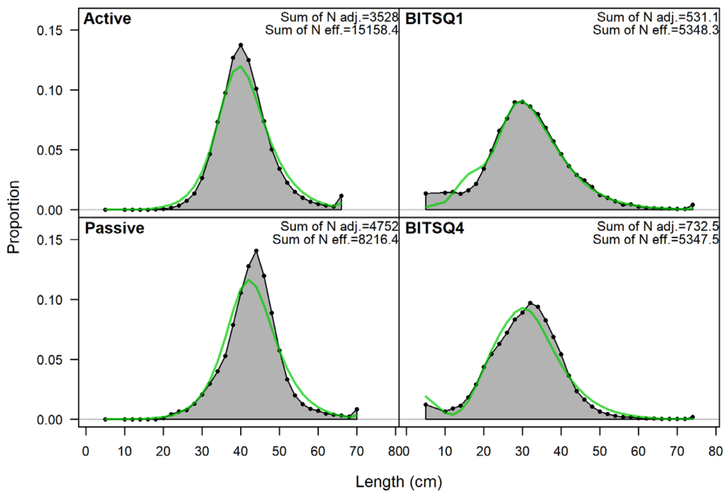


Figure 4.21. Model fits to length composition data (combined across years and quarters).

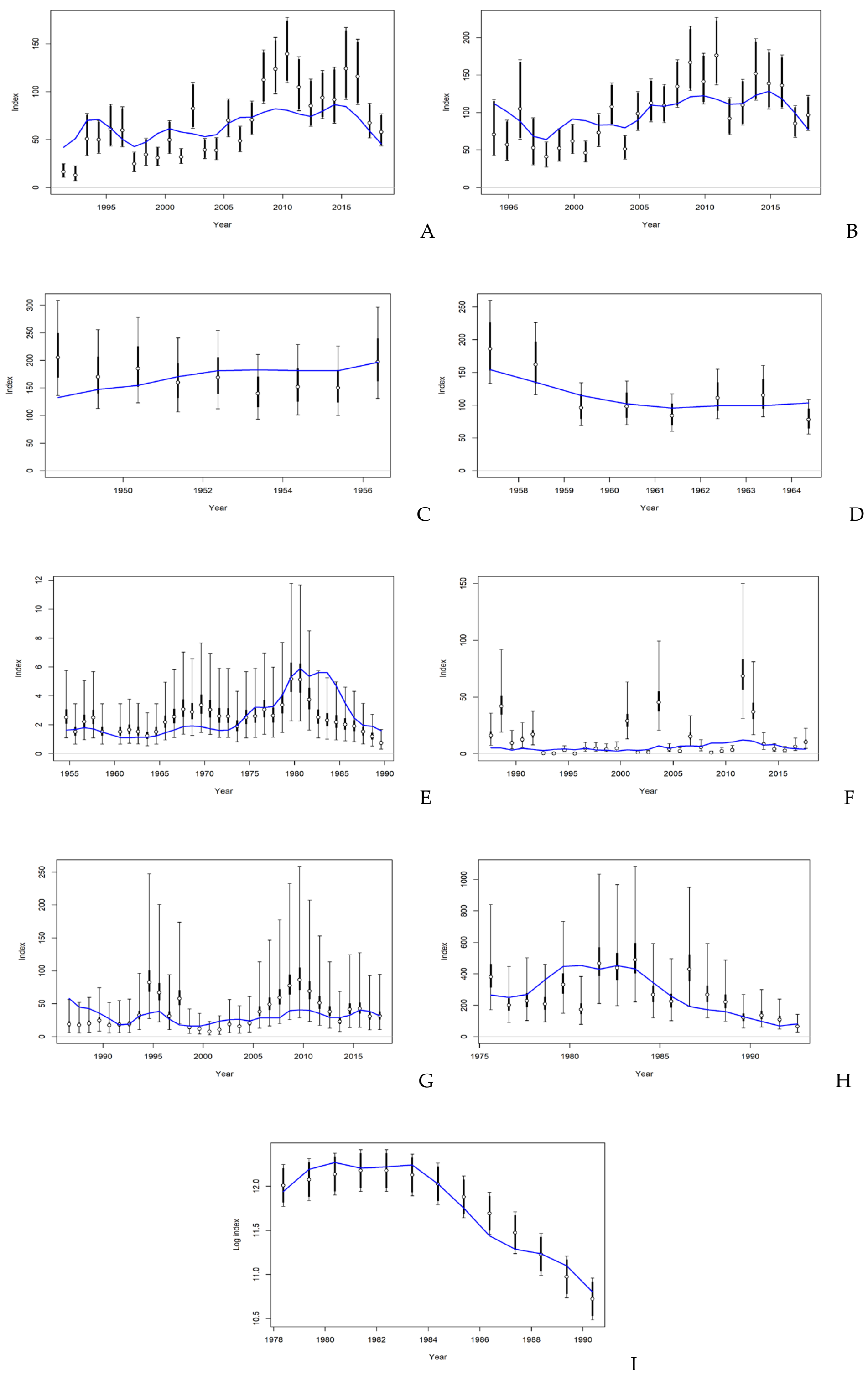


Figure 4.22. Model fits to different tuning indices. A- BITSQ1; B-BITSQ4; C- CommCpue1; D- CommCpue2; E- CommCpue3; F- Larvae; G- SSBEggProd; H- TrawlSurvey1; I- TrawlSurvey2 (see section 4.5.4 for description of the different indices).

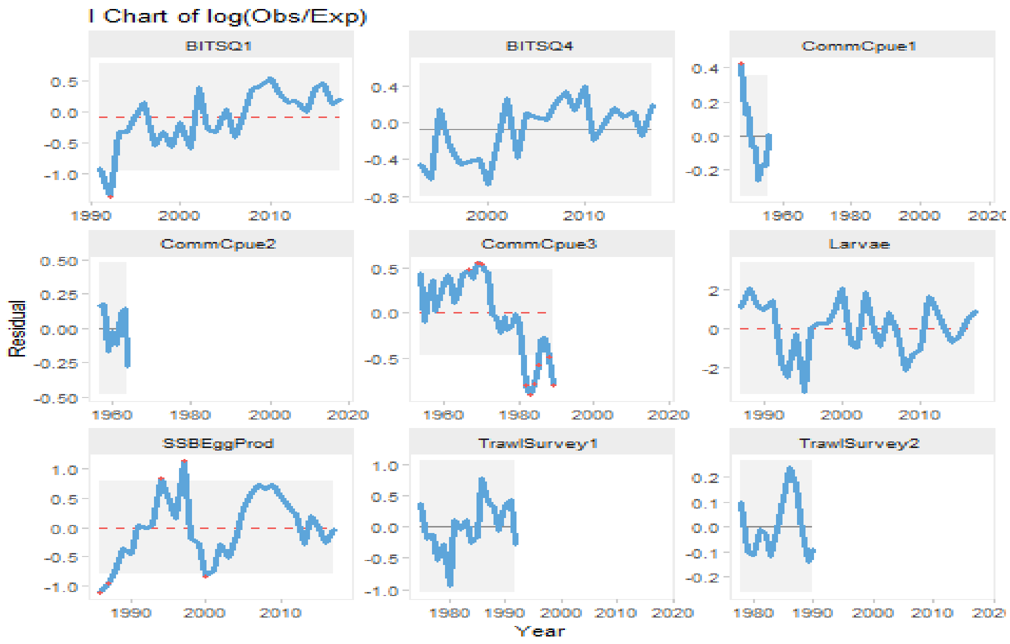


Figure 4.23. Residuals from runs test analyses, for the fit to survey indices.

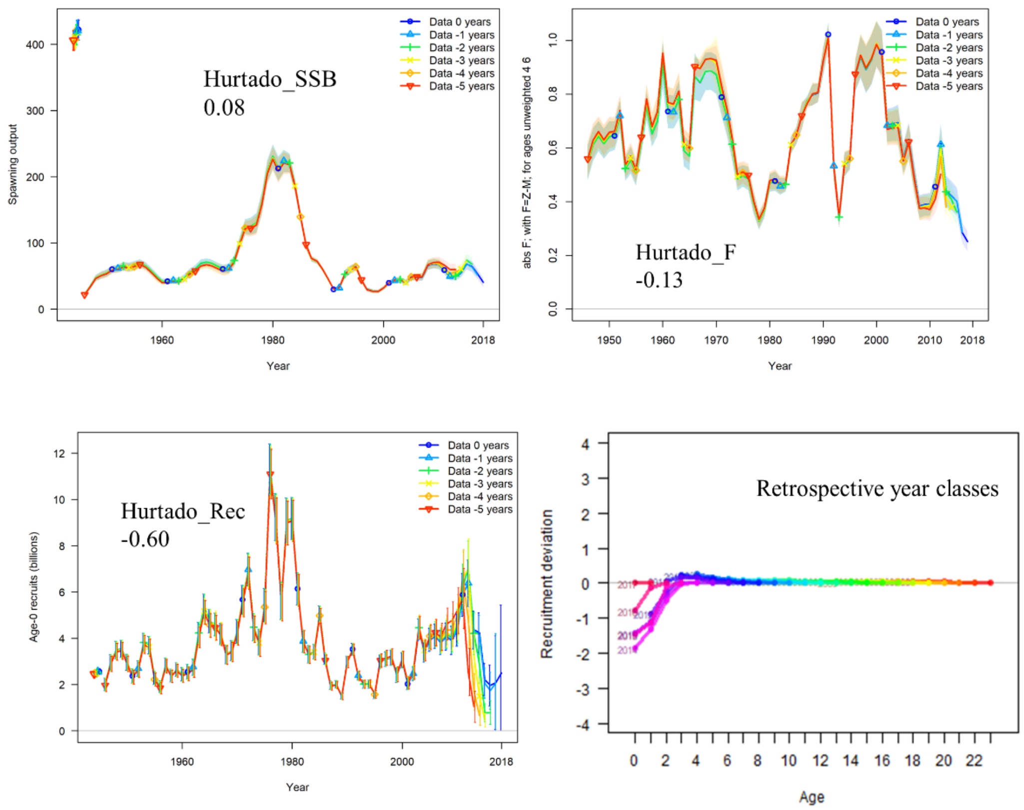


Figure 4.24 Retrospective analyses of the final Stock Synthesis model.

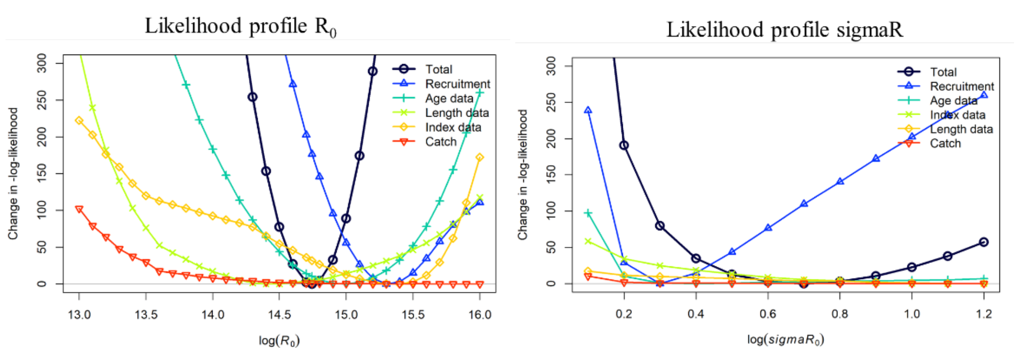


Figure 4.25. Likelihood profiles for R0 and sigmaR.

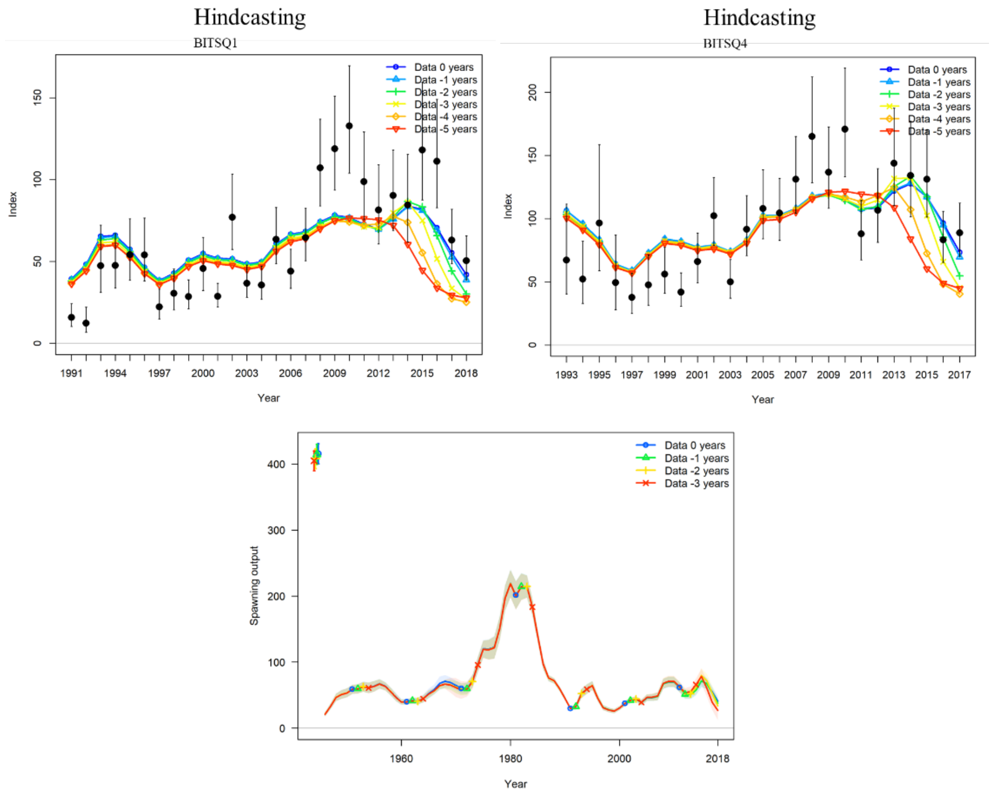


Figure 4.26. Results of hindcasting for BITS surveys, and SSB.

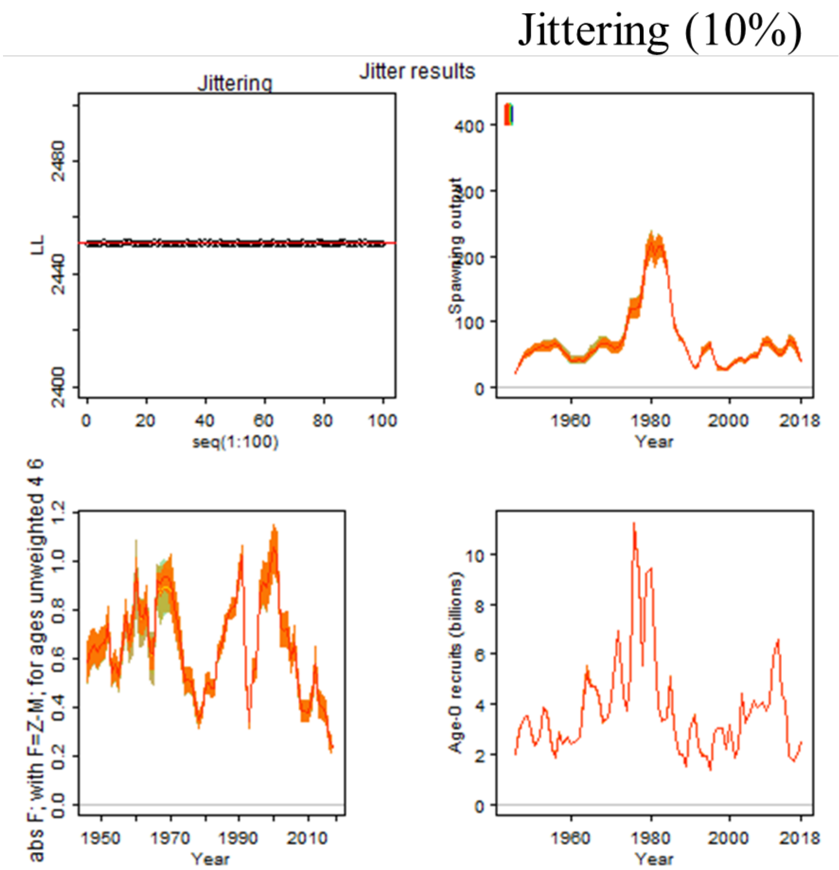


Figure 4.27. Results from jittering using 100 iterations and an average jitter of 10%.

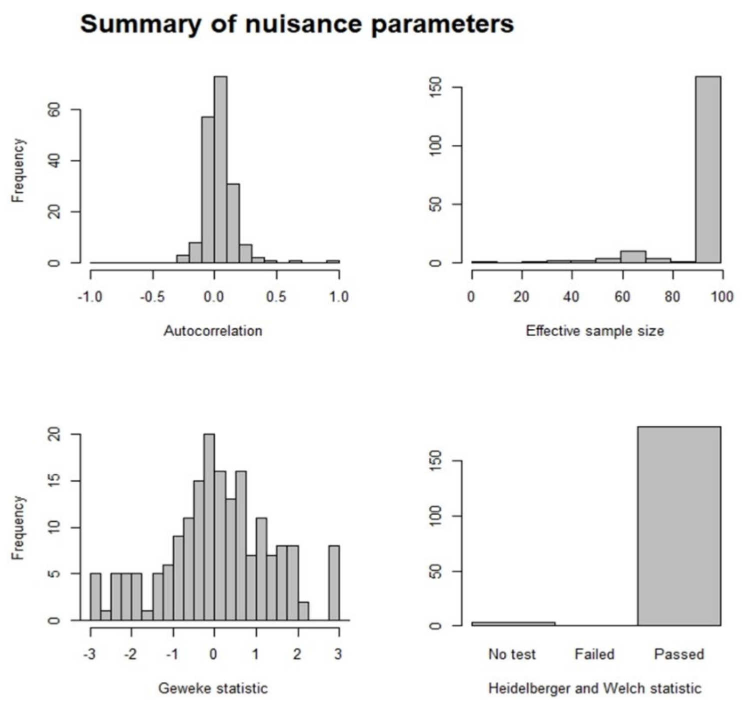


Figure 4.28. Statistics from the MCMC analysis.

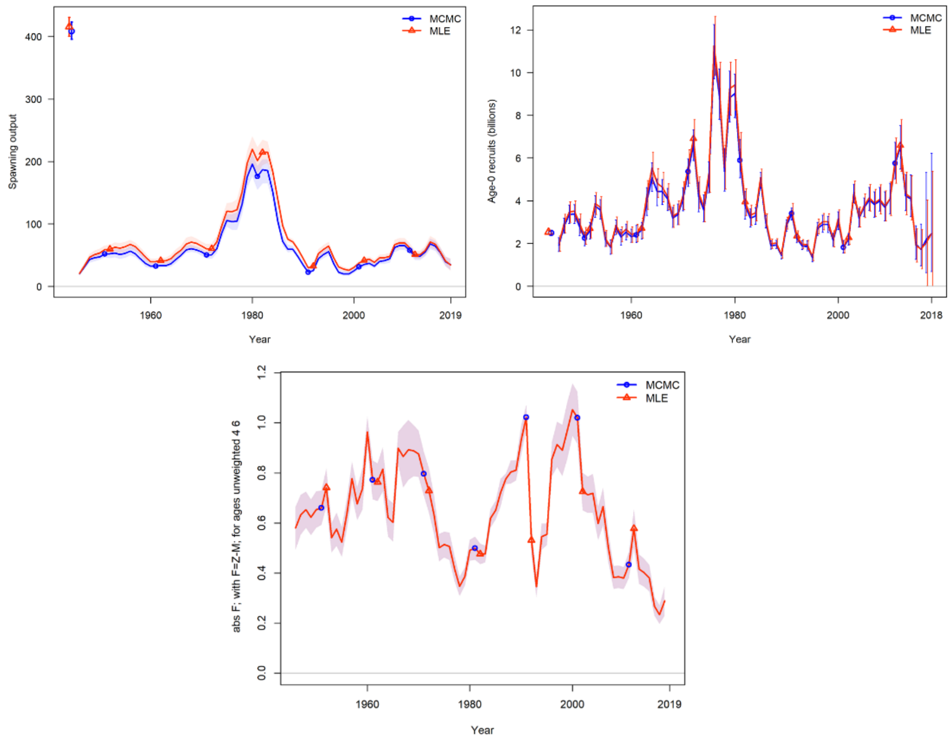


Figure 4.29. Results of the MCMC analysis in terms of SSB, R and F compared to MLE.

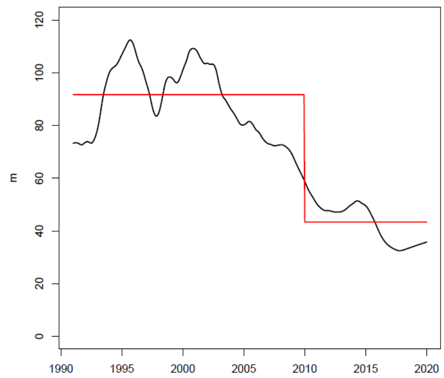
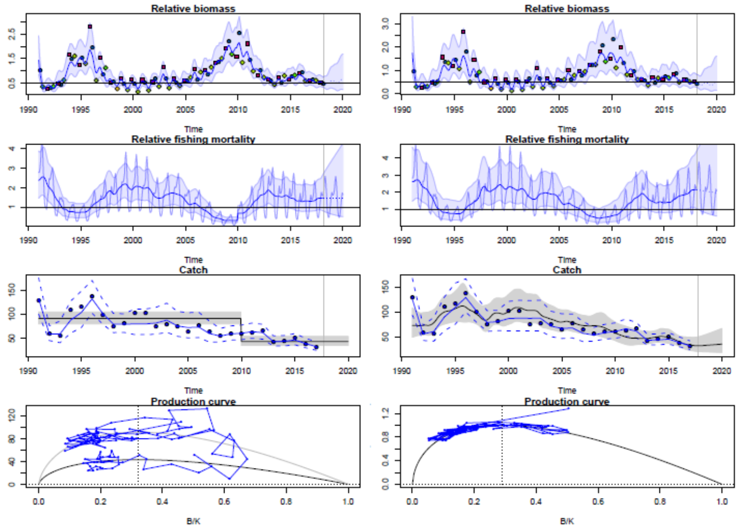


Figure 4.30. Change in productivity from SPICT model using different productivity regimes (red line) or a gradual change in productivity (black line).



Left: stepwise model, Right: gradual model

Figure 4.31. Results from SPICT model with stepwise and gradual change in productivity.

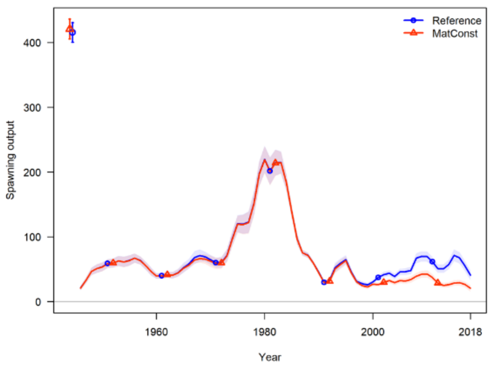


Figure 4.32. SSB (females only) taking into account the observed reduced size at maturation (reference run, blue line) compared to the biomass of the same size of cod that corresponded to the SSB before 2000s (L50 at 38 cm) (red line).

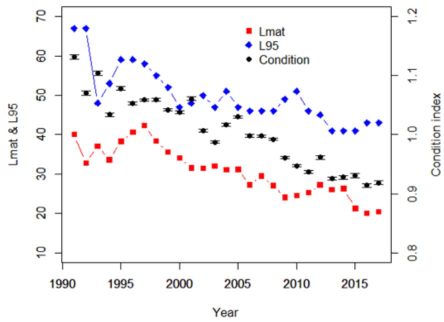


Figure 4.33. Changes in size at maturation (Lmat= L50), size structure of the stock (L95- length at 95th percentile of the length distribution) and nutritional condition of Eastern Baltic cod.

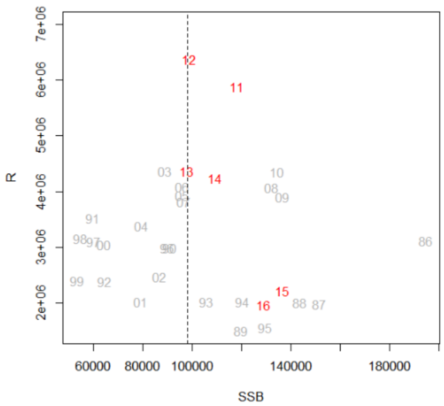


Figure 4.34. Stock–recruitment relationship, where the red points highlight the latest years and the latest strong year class from 2012 that was the basis for setting B_{lim}. Grey points corresponds to S-R in earlier years since 1986.

5 External Reviewers Comments

Meaghan Bryan (co-chair), Verena Trenkel, and Vladlena Gertseva acted as the external experts for the WKBALTCOD2 benchmark of Baltic cod. We evaluated the modelling methods used for the assessment of the Eastern and Western Baltic cod stocks on 4–8 February, 2019.

Assessment working documents and supporting materials were distributed prior to the benchmark meeting. Additionally, two WebEx meetings were organized prior to the in-person meeting. The first WebEx meeting provided an overview of the data evaluation process, the decisions made at the data evaluation workshop, and a general discussion about the proposed modelling frameworks for each stock. The second WebEx meeting focused on the preliminary Eastern Baltic cod (EBC) Stock Synthesis (SS) model. We appreciate that these materials were made available before the in-person benchmark meeting. One suggested improvement would be to provide a table summarizing the data sources reviewed at the data evaluation workshop, the decisions made about each, and the justification for the decision. Data sources that were not reviewed prior to the benchmark workshop, but intended to be used in the assessment model(s), and the corresponding issues/decisions should be included in the table. We also recommend these types of preparatory efforts continue in future as they improved the efficiency of the review process and the benchmark workshop.

We commend the workshop participants for their efforts during the benchmark process. The assessment team was asked to provide many additional analyses during the meeting. Their response to the requests was helpful in furthering our understanding of the assessment models and were successful in bringing useful information to the management process.

The reviewers confirm that the outcomes of the benchmark are appropriate to provide scientific advice.

The sections below summarize the discussions during the meeting and the recommendations made regarding these stocks.

Western Baltic cod

Issues addressed at the benchmark

A single modelling approach, the state space stock assessment model (SAM) was presented for the Western Baltic cod (WBC) stock assessment. The panel agreed this was a sensible model to provide management advice given its previous use for the WBC assessment and the overall model performance. Over the course of the week, alternative model runs were explored to better understand the model behaviour. The main issues that emerged were: 1) stock mixing in SD24; 2) how to best include the recreational fisheries data in the assessment model; 3) extending the assessment period back in time; 4) inclusion of the poundnet survey index in the assessment model; and 5) updating reference points.

Baltic cod are separated into two genetically different stocks, western and eastern. Although there is some evidence of mixing in a number of ICES subdivisions (SD), the predominant area of mixing is SD24. This mixing poses a challenge for splitting the landings and survey data. At the benchmark it was decided to split the survey data using a line at 13° East as attempts to use a more progressive split were not convincing due to lack of sufficient information on the stock origin of individuals to support such a model. This decision is practical and does not seem unduly influential for the stock assessment. The assumption underlying this decision is that the proportion of the WBC stock east of 13° is more or less constant in time. This assumption might or might not be violated as the large 2016 cohort ages. It is therefore recommended to monitor the spatial distribution of the 2016-year class.

Commercial landings were split between the two stocks using an annual splitting coefficient derived from otolith shape comparisons and some genetics data. Two estimation methods for deriving the splitting factor were presented at the meeting. Although they use the same data, some notable differences in the estimates occurred. This raises the fundamental question of the adequacy of the sampling design underlying the estimation of the splitting factor. Currently it is unknown how well the sampled individuals represent the mixed stocks targeted by the commercial fisheries. The panel recommends documenting where the sampled individuals come from in space and time, and evaluate how well the sample represents the fishery catches and possibly revise the sampling scheme. Further, based on the temporal stability of the currently available splitting coefficients it is recommended to consider simplifying the splitting method for the commercial landings, e.g. by using fixed coefficients or smoothed coefficients for time blocks. This would ensure that sampling variability has less impact on the assessment. Further, a scientific study should be conducted to confirm clearly formulated hypotheses about the ecological process underlying stock mixing (e.g. migration paths and times, habitat choice, return behaviour, etc.).

Recreational fisheries targeting Baltic cod occur in several countries. Recreational data from Germany were used in the previous WBC stock assessment. New estimates were presented for Denmark and Sweden. The estimates were obtained by combining information on removals with length samples in SD23 and age–length keys from SD22. In contrast to this, for commercial catches age–length keys from SD23 were used, which is somewhat inconsistent. The estimates indicate that the order of magnitude of removals by recreational fisheries and commercial fisheries is about the same in SD23. However, the overall contribution of recreational fisheries, at least in Denmark and Sweden, to WBC fishing mortality is much smaller than commercial fisheries and most likely independent of stock status. Thus, to remove sampling variability and as a practical way forward, it might be sufficient to assume constant recreational removals (in Denmark and Sweden) over a certain time frame. Indices of changes in recreational fisheries could then be monitored (e.g. number of tour boats) and a comprehensive study estimating detailed recreational catches be carried out only every couple of years, e.g. every five years or some other time interval. Given they are larger, German recreational catches will need to be monitored more closely using the current approach.

The panel agreed that the assessment model should start in 1985. The main reason for this was to include data encompassing a period with considerable contrast. Two periods of high and low spawning-stock biomass can be observed from 1985 to present. Older data exist, but are not currently included in the model. It would be interesting to try to include these data at a future benchmark to potentially provide improved MSY-based metrics.

The panel agreed that the abundance index derived from commercial poundnets be included in the model. This index represents a cooperative data collection program between stakeholders and scientists, covers shallower waters than the BITS, and provides information on age-0 and age-1 cod. Another advantage of the poundnet index is that it covers several months and hence smooths out environmental variations. It was decided to include the index for age-0 only, but it would be interesting to further explore the reliability of the age-1 index.

Ageing error, an issue raised as important to resolve during the WKBALTCOD 2015 benchmark, continues to be an issue. Results from a recent ageing inter-calibration study were presented at the meeting. They clearly indicate differences between countries in age reading. It is recommended to derive estimates of ageing error and bias that can be used to either internally inform the model or externally evaluate and better understand model residual patterns for the next benchmark.

The reference points were updated from the WKBALTCOD 2015 benchmark. There was some debate about which ICES stock–recruitment relationship category describes the trends (or lack thereof) for WBC. Most agreed it would best be defined as a hybrid of type-3 and type-5. Hence, in the absence of a clear continuous stock–recruitment relationship, it was proposed to use the average biomass of four recent years for which reasonably high recruitment occurred to estimate B_{lim}. The panel agreed this approach was sensible.

Eastern Baltic cod

Issues addressed at the benchmark

The Eastern Baltic stock is estimated to be in extremely depleted state. The historical fisheries in the 1980s resulted in the removal of larger fish. This in addition to changing environmental conditions have triggered changes in biological characteristics, including growth decline, earlier maturation, increase in natural mortality, and others. The biological parameters have been changing gradually and have posed a challenge in modelling the stock. A key issue addressed during this benchmark was how to best account for this changing productivity in the assessment model. Other keys issues discussed at the meeting included mixing of the Eastern and Western Baltic cod stocks in area SD 24, and approaches to split catch and survey data between the stocks in that area, and the incorporation of age–length keys (ALKs) in the assessment. In addition, model performance and variety of diagnostics were presented and evaluated.

The assessment uses the Stock Synthesis (SS) modelling platform (Methot and Wetzel, 2013). This is a flexible platform that allows one to incorporate a variety of available information, and explore various assumptions and modelling choices regarding the stock, including assumptions about time-varying parameters. Several modifications to the stochastic production model in continuous time (SPiCT), which has been used previously to classify stock status for EBC, were also presented as a potential alternative approach in case the SS based model was found to be inadequate to predict the dynamics of the stock.

A variety of SS model runs were reviewed at the meeting, including the proposed reference run and runs with alternative assumptions regarding mortality and growth parameters. In general, the model was able to fit the data relatively well and generate dynamics that are consistent with current knowledge of the stock. The values of estimated parameters were within those reported in the literature and fit the current knowledge of the stock. Several residual patterns emerged in the model fit to the data. The model was able to fit the general trend of the BITS indices; however, it underestimated the high index estimates in 2008–2011 and 2016–2017. The model also exhibited a systematic pattern in the fit to length composition data from the active and passive fleets and the BITS quarter 4 survey. At the meeting, additional model runs were completed to better understand the behaviour of the model and resolve these issues.

Several model runs were explored to improve the fit to the BITS indices of abundance. The model was run with the index time-series split in 2002 (when changes in gear and survey design occurred) and separate survey catchability parameters were estimated for each period. This was done to explore whether the statistical standardization of the BITS indices (done outside the model) was sufficient to account for changes in the survey that occurred. Two model runs with differing treatments of the BITS length composition data were also run. One run excluded the BITS length composition data for years prior to 2002 and the other reduced input sample sizes (i.e. these data were down-weighted) for years prior 2002. Both model runs were done to evaluate how earlier length compositions that have higher uncertainty impact the fit to survey index and length compositions. An improvement in the fit to the BITS indices was not achieved and most runs resulted in virtually identical estimates of stock dynamics. The run with indices being split in two parts and additional catchability parameters estimated produced unreasonable results, with regards to stock dynamics and biological parameter estimates.

To resolve the systematic pattern in the model fit to length composition data in the active and passive fleets and the BITS Q4 survey, the model was run with a more flexible double-normal selectivity curve. Using the double-normal relationship allows the selectivity for all the fleets to be dome-shaped (in the reference model, logistic selectivity curves are used). Selectivity was still estimated as asymptotic using the double-normal and these runs did not resolve the pattern.

The model was also run with alternative assumptions about ageing error to evaluate how ageing error assumptions affected the fit and model output, as imprecisions and bias in the age readings since 2000 have been a concern for the assessment of EBC. Stock synthesis is able to account for ageing error while estimating growth parameters and fitting the data. The ageing error matrices were not available for this benchmark; therefore, the model assumed that age estimates were precise. At the meeting, an ageing error matrix was developed that accounted only for imprecision in age estimates, and not the bias. The results of running the model with the alternative ageing error matrix did not change trends in the model output, and did not the resolve the residual pattern in model fit to length composition, but improved overall fit to commercial length composition data, fitting the peaks of the compositions better.

The panel agreed that although the reference model exhibited patterns in the residuals, overall the model generated stock dynamics that are consistent with the current knowledge of the stock and is an appropriate basis to provide management advice. The jitter diagnostics demonstrated that the model was stable and the retrospective pattern was minimal.

Eastern Baltic cod have experienced significant environmental and biological changes over the last 30 years. These changes include the loss of two important spawning grounds, declining body condition, reduced size at maturity, declining egg quality (e.g. poor buoyancy), among others due to historic fishing levels and the continuing and expanding hypoxic conditions in the Eastern Baltic, and increased parasitic infection rates. There was some debate about whether B_{lim} should be defined as B_{loss} given the relatively flat stock–recruitment relationship. B_{loss} was associated with SSB levels in the 1990s and the general thought was that this did not reflect the current conditions in the Eastern Baltic. More specifically, SSB is currently dominated by smaller fish that produce lower quality eggs. The suggested B_{lim} reference point for EBC corresponded to the lowest SSB observed that produced a recent (2012), strong year class. The panel agreed that this was a sensible decision given the changing environmental and biological conditions that EBC are currently experiencing. The changes this dynamic system is experiencing are expected to continue; therefore, the reference points should be considered for a short-term (2-3 years) period.

Prioritized list of recommendations for future work

Ageing error and stock mixing/splitting in area SD24 are two important issues that are common between the assessments and efforts to address them should be continued. The approaches for splitting the survey and commercial catch data in area SD24 were slightly different. The survey data were split at 13° East, whereas, the commercial catch data are split according to proportional constant based on genetics data. Efforts should be made to make these decisions consistent between the survey and commercial catch data. As a first step, it may be prudent to conduct comparable sensitivity runs for both stocks using a common set of assumptions to determine the impact the splitting assumptions have on the two assessments. This would help to better understand the importance of the splitting assumptions on the assessment results and to justify the directed effort and funds needed to adequately address this issue (see the WBC section for detailed research recommendations).

Ageing error is an issue that should be addressed for both stocks. Some of the residual patterns in both stock assessments were likely due to assuming the ages were precisely known. The EBC stock assessment model was developed in SS and can internally accommodate an ageing error matrix that accounts for precision and bias. An ageing error matrix was applied in one of the alternative model runs for EBC and did lead to an improvement to the overall fit to the length composition data. The ageing error estimates, however, were not fully evaluated prior to the meeting and did not account for bias in the age readings. Given that this is a critical issue for both stock assessments, it is of utmost importance that ageing error and bias statistics be developed and thoroughly reviewed at the data evaluation workshop prior to the next benchmark.

The EBC stock assessment model accounts for a number of time-varying processes. Given the level of complexity, the panel suggests that a series of bridging models be developed. This would entail the stock assessors to start with a simple model (i.e. no time-varying processes) and add complexity sequentially. This will help to evaluate the magnitude of the impact of the individual assumptions about time-varying growth, natural mortality, length-at-maturity, etc., on the assessment results.

It became clear through the discussions during the benchmark workshop that the survey design of the BITS has changed over time. A single alternative run for EBC where the BITS index was split by estimating catchability parameters for two different periods was evaluated. This was done to improve the fit to the index; however, the residual pattern was unchanged. The residual pattern suggested the potential for non-stationarity in the index; hence, the panels recommends evaluating whether it would be justified to assume that catchability has changed over time for the next benchmark assessment of EBC.

A table summarizing the data sources reviewed at the data evaluation workshop, the decisions made about each, and the justification for the decision should be provided to the external experts prior to or at the beginning of the next benchmark. Additionally, data sources that were not reviewed prior to the benchmark workshop and the corresponding unresolved issues/decisions should be included in the table.

6 New Stock annexes

The new stock Annexes are available here:

Western Baltic cod in Subdivisions 22-24
http://ices.dk/sites/pub/Publication%20Reports/Stock%20Annexes/2019/cod.27.22-24_SA.pdf

Cod (*Gadus morhua*) in Subdivisions 24–32, eastern Baltic stock
http://ices.dk/sites/pub/Publication%20Reports/Stock%20Annexes/2019/cod.27.24-32_SA.pdf

7 References

Bélanger, J., Gagnon, D. 1993. Classification and proportion estimation. Aust. J. Stat. 35, 19–28. DOI 10.1111/j.1467-842X.1993.tb01309.x.

Berg, Casper W. and Kasper Kristensen. 2012. Spatial age–length key modelling using continuation ratio logits. Fisheries Research , 129:119-126.

Birjukov, N. P. 1970. Baltijskaja treska [Baltic cod]. AtlantNIRO, Kaliningrad. 165 pp. (in Russian).

Bolles, K.L., Begg, G.A., 2000. Distinctions between silver hake (*Merluccius bilinearis*) stocks in U.S. waters of the northwest Atlantic based on whole otolith morphometric. Fish. Bull. 98, 451–462.

Campana, S.E., Casselman, J.M., 1993. Stock discrimination using otolith shape analysis. Can. J. Fish. Aquat. Sci. 50, 1062–1083.

Cardinale, M., Doering-Arjes, P., Kastowsky, M., Mosegaard, H., 2004. Effects of sex, stock, and environment on the shape of known-age Atlantic cod (*Gadus morhua*) otoliths. Can. J. Fish. Aquat. Sci. 61, 158–167.

Casini M., Käll, F., Hansson, M., Plikshs, M., Baranova, T., Karlsson, O., Lundström, K., Neuenfeldt, S., Gårdmark, A., Hjelm, J. (2016). Hypoxic areas, density-dependence and food limitation drive the body condition of a heavily exploited marine fish predator. Royal Society Open Science. 3: 160416.

Capizzano, C. W., Mandelman, J. W., Hoffman, W. S., Dean, M. J., Zemeckis, D. R., Benoît, H. P., Kneebone, J., *et al.* 2016. Estimating and mitigating the discard mortality of Atlantic cod (Gadus morhua) in the Gulf of Maine recreational rod-and-reel fishery. ICES Journal of Marine Science: Journal du Conseil, 73: 2342-2355.

Chhikara, R. S. 1986. Error analysis of crop acreage estimation using satellite data. Technometrics 28, 73–80.

Dementjeva, T. F. 1959. Some data on the life history and fishery of cod in the central Baltic. Rapports et Procès-Verbaux des Réunions du Conseil Permanent International pour l'Exploration de la Mer, 147: 68–73.

Eero, M., Köster, F.W., Plikshs, M., Thurow, F. 2007. Eastern Baltic cod (*Gadus morhua callarias*) stock dynamics: Extending the analytical assessment back to the mid-1940s. *ICES Journal of Marine Science*. 64: 1257–1271.

Eero, M., Köster, F.W., MacKenzie, B.R. 2008. Reconstructing historical stock development of Atlantic cod (*Gadus morhua*) in the eastern Baltic Sea before the beginning of intensive exploitation. *Can. J. Fish. Aquat. Sci*. 65: 2728–2741

Eero, M., Vinther, M., Haslob, H., Huwer, B., Casini, M., StorrPaulsen, M., and Köster, F. W. 2012. Spatial management of marine resources can enhance the recovery of predators and avoid local depletion of forage fish. Conservation Letters 5: 486–492.

Eero, M., Hemmer-Hansen, J., Hüssy, K., 2014. Implications of stock recovery for a neighbouring management unit: experience from the Baltic cod. ICES J. Mar. Sci. 71, 1458-1466.

Eero, M., Hjelm, J., Behrens, J., Buckmann, K., Casini, M., Gasyukov, P., Horbowy, J., Hüssy, K., Kirkegaard, E., Kornilovs, G., *et al*. 2015. Eastern Baltic cod in distress: an ecological puzzle hampering scientific guidance for fisheries management. ICES Journal of Marine Science 72: 2180–2186.

Ferter, K., Weltersbach, M. S., Humborstad, O. B., Fjelldal, P. G., Sambraus, F., Strehlow, H. V., and Volstad, J. H. 2015. Dive to survive: effects of capture depth on barotrauma and post-release survival of Atlantic cod (Gadus morhua) in recreational fisheries. ICES Journal of Marine Science, 72: 2467-2481.

Ferter, K., Weltersbach, M. S., Strehlow, H. V., Vølstad, J. H., Alós, J., Arlinghaus, R., Armstrong, M., *et al*. 2013. Unexpectedly high catch-and-release rates in European marine recreational fisheries: implications for science and management. ICES Journal of Marine Science: Journal du Conseil, 70: 1319-1329.

Francis, R.I.C. 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68: 1124–1138.

Freitas, C., Olsen, E.M., Knutsen, H., Albretsen, J., Moland, E., 2016. Temperature-associated habitat selection in a cold-water marine fish. J. Anim. Ecol. 85, 628–637. https://doi.org/10.1111/1365-2656.12458.

Gibbons, J.D., Chakraborti, S., 1992. Nonparametric Statistical Inference, thirdedition, Marcel Dekker, Inc., New York.

Healy, J. D. 1981. The effects of misclassification error on the estimation of several population proportions. The Bell Syst. Tech. J. 60, 697–705. DOI 10.1002/j.1538-7305.1981.tb00257.x.

Hemmer-Hansen J, Hüssy K, Baktoft , H, Huwer, B., Bekkevold, D., Haslob, H., Herrmann, J-P,. Hinrichsen, H-H., Krumme, U, Mosegaard, H., Nielsen, E E., Reusch, T. B. H., Storr-Paulsen, M., Velasco, A., Dewitz, B. v., Dierking, J. Eero, M. Genetic analyses reveal complex dynamics within a marine fish management area. Evol Appl. 2019;00:1–15. https://doi.org/10.1111/eva.12760

Hinrichsen, H. H., von Dewitz, B., Dierking, J., Haslob, H., Makarchouk, A., Petereit, C., and Voss R. 2016. Oxygen depletion in coastal seas and the effective spawning-stock biomass of an exploited fish species. Royal Society Open Science, 3: 150338.

Horbowy, J., Podolska, M., Nadolna-Ałtyn, K. 2016. Increasing occurrence of anisakid nematodes in the liver of cod (Gadusmorhua) from the Baltic Sea: Does infection affect the condition and mortality of fish? Fisheries Research 179: 98–103

Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J.,Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L.,Ono, K., Vert-Pre, K.A., Whitten, A.R., Punt, A.E., 2014. Looking in the rear-viewmirror: bias and retrospective patterns in integrated, age-structured stockassessment models. ICES J. Mar. Sci. 72, 99–110.

Hüssy, K., Mosegaard,H., Albertsen, C.M., Nielsen, E.E., Hemmer-Hansen,J., Eero, M. 2016a. Evaluation of otolith shape as a tool for stock discrimination in marine fishes using Baltic Sea cod as a case study. Fisheries Research, 174: 210-218.

Hüssy, K., Hinrichsen, H.-H., Eero, M., Mosegaard, H., Hemmer-Hansen, J., Lehmann, A., and Lundgaard, L. S. 2016b. Spatio-temporal trends in stock mixing of eastern and western Baltic cod in the Arkona Basin and the implications for recruitment. ICES Journal of Marine Science, 73 (2): 293-303 , doi: 10.1093/icesjms/fsv227.

Hüssy, K. 2011. Review of western Baltic cod (*Gadus morhua*) recruitment dynamics. ICES Journal of Marine Science, 68: 1459–1471.

Hüssy, K., Hinrichsen, H.-H., and Huwer, B. 2012. Hydrographic influence on the spawning habitat suitability of western Baltic cod (*Gadus morhua*) – ICES Journal of Marine Science, 69: 1736–1743.

Hyndman, R. J., Athanasopoulos, G. (2013) Forecasting: principles and practice. OTexts.org/fpp/.

ICES WGBFAS 2018. Report of the Baltic Fisheries Assessment Working Group (WGBFAS), 6–13 April 2018, ICES Headquarters, Copenhagen, Denmark. ICES CM 2018/ACOM:11.ICES.

ICES WKBALT 2013. Report of the Benchmark Workshop on Baltic Multispecies Assessments (WKBALT 2013), 4–8 February 2013, Copenhagen, Denmark. ICES CM 2013/ACOM:43.

ICES 2017. Report of the Workshop on Biological Input to Eastern Baltic Cod Assessment (WKBEBCA). ICES CM 2017/SSGEPD:19

ICES, 2017. Manual for the Baltic International Trawl Surveys (BITS). Ser. ICES Surv. Protoc. SISP 7 - BITS 95. https://doi.org/10.17895/ices.pub.2883.

ICES 2018. Workshop on evaluation of Input Data to Eastern Baltic Cod stock Assessment (WKIDEBCA). 23-25 January 2018, Copenhagen, Denmark

ICES 2014a. Report of the Workshop on Scoping for Integrated Baltic Cod Assessment (WKSIBCA). ICES CM 2014/ACOM:62.

ISSF, 2011. Report of the 2011 ISSF Stock Assessment Workshop Rome, Italy, March14-

17, 2011. ISSF Technical Report 2011-02. International Seafood Sustainability Foundation, Washington, D.C., USA.

Kell, L. *et al.* 2016. Evaluation of the prediction skill of stock assessment using hindcasting. Fisheries Research, 183: 119-127.

Kristensen, K. and Casper W. Berg. Datras package for r.http://rforge.net/DATRAS/,2012.

Köster F. W., Huwer, B, Hinrichsen, HH, Neumann, V, Makarchouk, A Eero, M, Tomkiewicz, J, Margonski, P , Temming, A, Hermann, JP , Oesterwind, D, Dierking, J, Kotterba, P, Plikshs, M. 2017. Eastern Baltic cod recruitment revisited-dynamics and impacting factors. ICES Journal of Marine Science, 74;1 page 3:19.

Kolody *et al.* 2019, Recruitment in tuna RFMO stock assessment and management: A review of current approaches and challenges.doi.org/10.1016/j.fishres.2018.11.031.

Kristensen, K. and Casper W. Berg. Datras package for r.http://rforge.net/DATRAS/,2012.

Lablaika, I., Hoziosky, S., Jushkevitz, Z. 1991. The USSR Baltic cod catches and relevant fishing effort dynamics in the ICES Subdivisions 26 and 28. ICES C.M.1991/J:22

Lass, H.U., Mohrholz, V., 2003. On dynamics and mixing of inflowing saltwater in the Arkona Sea. J. Geophys. Res. 108(C2), 3042, doi:10.1029/2002JC001465.Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology, 49: 627–642.

Matthäus, W., and Franck, H. 1992. Characteristics of major Baltic inflows – a statistical analysis. Continental Shelf Research, 12: 1375–1400.Mehrdana, F., Bahlool, Q. Z., Skov, J., Marana, M. H., Sindberg, D., Mundeling, M.,Overgaard,B.C.,etal.2014.Occurrenceofzoonotic nematodes Pseudoterranovadecipiens, Contracaecum osculatum and Anisakis simplex in cod (Gadus morhua) from the Baltic Sea. Veterinary Parasitology, 205: 581–587.

McQueen, K., Hrabowski, J., Krumme, U., 2018. Age validation of juvenile cod in the western Baltic Sea. ICES Journal of Marine Science. DOI:10.1093/icesjms/fsy175Methot, R.D., Wetzel, C.R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142 (2013) 86–99.

Mérigot, B., Letourneur, Y., Lecomte-Finiger, R., 2007. Characterization of local populations of the common sole *Solea solea* (*Pisces, Soleidae*) in the NW Mediterranean through otolith morphometrics and shape analysis. Mar. Biol. 151, 997–1008.

Orio, A., Florin, A.-B., Bergstro¨m, U., Sics, I., Baranova, T., and Casini, M. 2017. Modelling indices of abundance and size-based indicators of cod and flounder stocks in the Baltic Sea using newly standardized trawl survey data. – ICES Journal of Marine Science, doi:10.1093/icesjms/ fsx005.

Paul, K., Oeberst, R., Hammer, C., 2013. Evaluation of otolith shape analysis as a tool for discriminating adults of Baltic cod stocks. J. Appl. Ichthyol. 29, 743-750.

Pihl, L., Ulmestrand, M., 1993. Migration pattern of juvenile cod (Gadus morhua) on the Swedish west coast. ICES J. Mar. Sci. 50, 63–70.

Punt, A.E., 2017. Some insights into data weighting in integrated stock assessments. Fish. Res. 192, 52–65. http://dx.doi.org/10.1016/j.fishres.2015.12.006.

Reeves, S. A. 2003. A simulation study of the implications of age-reading errors for stock assement and management advice. – ICES Journal of Marine Science, 60: 314–328.

SEDAR 40, 2015. Atlantic Menhaden Stock Assessment Report. SEDAR, NorthCharleston, SC643.

Schinke, H., and Matthäus, W. 1998. On the causes of major Baltic inflows – an analysis of long time series. Continental Shelf Research, 18: 67–97.

Scientific, Technical and Economic Committee for Fisheries (STECF) – Fisheries Dependent- Information – Classic (STECF-17-09). Publications Office of the European Union, Luxembourg 2017, ISBN 978-92-79-67481-5, doi:10.2760/561459, JRC107598.

Sokolova, Maria and Buchmann, Kurt and Huwer, Bastian and Kania, PW and Krumme, U and Galatius, Anders and Hemmer-Hansen, J and Behrens, Jane. (2018). Spatial patterns in infection of cod Gadus morhua with the seal-associated liver worm Contracaecum osculatum sensu stricto from the Skagerrak to the central Baltic Sea. Marine Ecology Progress Series. 606. 10.3354/meps12773.

Strehlow, H. V., Schultz, N., Zimmermann, C., and Hammer, C. 2012. Cod catches taken by the German recreational fishery in the western Baltic Sea, 2005 - ì2010: implications for stock assessment and management. ICES Journal of Marine Science: Journal du Conseil, 69: 1769–1780.

Subbey, S., 2018. Parameter estimation in stock assessment modelling: caveats with gradient based algorithm. ICES Journal of Marine Science, doi:10,1093/icesjms/fsy044.

Then, A.Y., J.M. Hoenig, N.G. Hall and D.A. Hewitt. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES Journal of Marine Science 72 (1): 82-92. doi: 10.1093/icesjms/fsu136.

Thurow, F., and Weber, W. 1992. Catch rates versus biomass in Baltic cod. ICES CM 1992/J:37.

Valentinsson D., Ringdahl K., Storr-Paulsen M., Madsen N. (2019) The Baltic Cod Trawl Fishery: The Perfect Fishery for a Successful Implementation of the Landing Obligation?. In: Uhlmann S.S., Ulrich C., Kennelly S.J. (eds) The European Landing Obligation. Springer, Cham

Weltersbach, M. S., and Strehlow, H. V. 2013. Dead or alive—estimating post-release mortality of Atlantic cod in the recreational fishery. ICES Journal of Marine Science: Journal du Conseil, 70: 864–872.

Weltersbach, M.S., Lewin, W.-C., Gröger, J.P., Strehlow, H.V. 2019. Effect of lure and bait type on catch, size, hooking location, injury and bycatch in the western Baltic Sea recreational cod fishery. Fisheries Research, 210: 121–130.Wilkins, M. F., Boddy, L., Morris, C. W. 2000. Proportion estimation with confidence limits. J. Microbiol. Methods 43, 55–64. DOI 10.1016/S0167-7012(00)00204-9.

Wilkins, M. F., Boddy, L., Morris, C. W. 2000. Proportion estimation with confidence limits. J. Microbiol. Methods 43, 55–64. DOI 10.1016/S0167-7012(00)00204-9.

Annex 1: List of participants

Name	Institute	Email
Alessandro Orio	Department of Aquatic Resources-SLU Aqua, Sweden	alessandro.orio@slu.se
Anastasiia Karpushevskaia	AtlantNIRO, Russia	anastasia0006@mail.ru
Anders Nielsen	DTU Aqua -National Institute of Aquatic Resources, Denmark	an@aqua.dtu.dk
Andreas Sundelöf	Department of Aquatic Resources-SLU Aqua, Sweden	andreas.sundelof@slu.se
Casper Willestofte Berg	DTU Aqua -National Institute of Aquatic Resources, Denmark	cbe@aqua.dtu.dk
Christoffer Moesgaard Albertsen	DTU Aqua -National Institute of Aquatic Resources, Demark	cmoe@aqua.dtu.dk
Conrad Stralka	Baltic Sea 2020, Sweden	conrad.stralka@balticsea2020.org
Francesca Vitale	Department of Aquatic Resources-SLU Aqua, Sweden	francesca.vitale@slu.se
Franziska Schade	Thünen-Institute of Baltic Sea Fisheries, Germany	franziska.schade@thuenen.de
Fritz W. Köster	DTU Aqua -National Institute of Aquatic Resources, Demark	fwk@aqua.dtu.dk
Hans Jakob Olesen	DTU Aqua -National Institute of Aquatic Resources, Denmark	hjo@aqua.dtu.dk
Harry Strehlow	Thünen-Institute of Baltic Sea Fisheries, Germany	harry.strehlow@thuenen.de
Hege Sande	Department of Aquatic Resources-SLU Aqua, Sweden	hege.sande@slu.se
Henrik Mosegaard	DTU Aqua -National Institute of Aquatic Resources, Denmark	hm@aqua.dtu.dk
Jan Horbowy	National Marine Fisheries Research Institute, Poland	horbowy@mir.gdynia.pl

Name	Institute	Email
Jane Behrens	DTU Aqua -National Institute of Aquatic Resources, Denmark	jabeh@aqua.dtu.dk
Joakim Hjelm	Department of Aquatic Resources-SLU Aqua, Sweden	joakim.hjelm@slu.se
Johan Lövgren	Department of Aquatic Resources-SLU Aqua, Sweden	johan.lovgren@slu.se
Jonna Tomkiewikz	DTU Aqua -National Institute of Aquatic Resources, Denmark	jt@aqua.dtu.dk
Karin Hüssy	DTU Aqua -National Institute of Aquatic Resources, Denmark	kh@aqua.dtu.dk
Kate McQueen	Thünen-Institute of Baltic Sea Fisheries, Germany	kate.mcqueen@thuenen.de
Keith Brander	DTU Aqua -National Institute of Aquatic Resources, Denmark	kbr@aqua.dtu.dk
Kåre Nolde Nielsen	University of Tromsø, Norway	kare.nolde.nielsen@uit.no
Marc Simon Weltersbach	Thünen-Institute of Baltic Sea Fisheries, Germany	simon.weltersbach@thuenen.de
Marcin Ruciński	Low Impact Fishers of Europe, UK	bans@lifeplatform.eu
Margit Eero	DTU Aqua -National Institute of Aquatic Resources, Denmark	mee@aqua.dtu.dk
Maria Pierce	Thünen-Institute of Baltic Sea Fisheries, Germany	maria.pierce@thuenen.de
Marie Plambech	DTU Aqua -National Institute of Aquatic Resources, Denmark	mpla@aqua.dtu.dk
Marie Storr-Paulsen	DTU Aqua -National Institute of Aquatic Resources, Denmark	msp@aqua.dtu.dk
Maris Plikshs	Institute of Food Safety Animal Health and Environment (BIOR), Latvia	Maris.Plikss@bior.lv
Massimiliano Cardinale	Department of Aquatic Resources-SLU Aqua, Sweden	massimiliano.cardinale@slu.se
Meaghan Bryan (Chair)	National Oceanic and Atmospheric Administration, USA	Meaghan.bryan@noaa.gov

Name	Institute	Email
Michael Andersen	BSAC and Danish Fishermen PO, Denmark	ma@dkfisk.dk
Michele Casini (Chair)	Department of Aquatic Resources-SLU Aqua, Sweden	michele.casini@slu.se
Mikaela Bergenius	Department of Aquatic Resources-SLU Aqua, Sweden	mikaela.bergenius@slu.se
Monica Mion	Department of Aquatic Resources-SLU Aqua, Sweden	monica.mion@slu.se
Nils Höglund	BSAC and Coalition Clean, Sweden Baltic	nils.hoglund@ccb.se
Sebastian Linke	University of Gothenburg, Sweden	sebastian.linke@gu.se
Sofia Carlshamre	Department of Aquatic Resources-SLU Aqua, Sweden	sofia.carlshamre@slu.se
Staffan Larsson	Swedish Cod Fishermen's Producer Organisation, Sweden	staffan@sfpo.se
Stefanie Haase	Thünen-Institute of Baltic Sea Fisheries, Germany	stefanie.haase@thuenen.de
Uwe Krumme	Thünen-Institute of Baltic Sea Fisheries, Germany	uwe.krumme@thuenen.de
Verena Trenkel	Ifremer, France	Verena.Trenkel@ifremer.fr
Vladlena Gertseva	National Oceanic and Atmospheric Administration, USA	vladlena.gertseva@noaa.org
Viktoriia Amosova	AtlantNIRO, Russia	amosova@atlantniro.ru
Xochitl Cormon	Institute for Marine Ecosystem and Fishery Science, Hamburg University and Biodiversity Economics, German Centre for Integrative Biodiversity Research (iDiv), Germany	xochitl.cormon@uni-hamburg.de
Yvette Heimbrand	Department of Aquatic Resources-SLU Aqua, Sweden	yvette.heimbrand@slu.se

Annex 2: Agenda

Benchmark Workshop on Baltic Cod Stocks (WKBALTCOD2)

Date & Location: 4-8 February, 2019, ICES HQ, Copenhagen, Denmark

Start: Monday 04/02, 09.00. **Finish:** Friday 08/02, 13.00

<table>
<tr><th></th><th>Eastern Baltic cod</th><th>Western Baltic cod</th></tr>
<tr><td>4/2</td><td colspan="2">FOLLOW UP ON DATA ISSUES FROM THE DATA MEETING:
1. Summary of EBC and WBC, from the data preparation meeting and intersessional work (Margit and Rie)
2. BITS indices, including stock mixing (Casper)</td></tr>
<tr><td></td><td>INTRODUCTION OF ASSESSMENT MODELS:
• SPICT (Casper)
• SS reference run, including overview of the input data used (Max/Margit)</td><td>• Recreational catch data</td></tr>
<tr><td>5/2</td><td colspan="2">ASSESSMENT RUNS:</td></tr>
<tr><td></td><td>• Additional SS runs compared to reference run (Max/Margit)
1. Time for preparing any additional runs if needed</td><td>2. Exploratory SAM runs</td></tr>
<tr><td>6/2</td><td colspan="2">FINAL ASSESSMENTS, INTRO to FORECAST & REFERENCE POINTS:</td></tr>
<tr><td></td><td>• Defining final configuration/input for the final model(s)
• S-R relationships and possible Blim for EB cod (Margit)</td><td>1. Method to be used for splitting commercial catch btw stocks in future WGBFAS (Christoffer)
2. Final assessment with SAM and forecast settings</td></tr>
<tr><td>7/2</td><td colspan="2">FORECAST /REFERENCE POINTS/ APPROACH TO ADVICE:
• Possible approach to forecast/advice based on SPICT (Casper)
• Possible FMSY for EB cod from Eqsim (Max)
• Possible alternative approach(es) to EB cod forecast/advice (Max)

• S-R relationships, possible MSY ref points for WB cod (Marie)
• Time for preparing anything additional concerning ref points if needed</td></tr>
<tr><td>8/2</td><td colspan="2">FINAL CONCLUSIONS AND WRAP-UP:
• Final decisions on reference points/forecast/advice approach for EB and WB cod
• Any other remaining decisions</td></tr>
</table>

Annex 3: Resolution

WKBALTCOD2 – Benchmark Workshop on Baltic Cod Stocks

A **Benchmark Workshop on Baltic Cod Stocks** (WKBALTCOD), attended by invited external experts Vladlena Gersteva, USA, and Verena Trenkel, France, will be established and will meet in Copenhagen, Denmark, 15–19 October 2018 for a data evaluation meeting (chaired by ICES Chair Johan Lövgren) and in Copenhagen, Denmark 4–8 February 2019 for a Benchmark meeting (chaired by External Chair Meaghan Bryan, USA, and ICES Chair Michele Casini, Sweden) 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:

   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. If a category 1 assessment method can not be agreed, then an alternative method (the former method, or following the ICES data-limited stock approach) should be put forward as a basis for the assessment and advice;

c) Re-examine and update (if necessary) MSY and PA reference points according to ICES guidelines (see Technical document on reference points);

d) Prioritize recommendations for future improving of the assessment methodology and data collection;

e) As part of the evaluation:

   i) Conduct a 5-day data evaluation workshop (DEWK). Stakeholders are invited to contribute data (including data from non-traditional sources) and to contribute to data preparation and evaluation of data quality. As part of the data evaluation workshop consider the quality and compiling methodology for all input data for stock assessment, including catch data. For both stocks, produce working documents at least 7 days prior to the meeting, describing the input data intended to be used in stock assessment, to be discussed, reviewed and approved during DEWK.

   ii) Following the DEWK, produce working documents to be reviewed during the Benchmark meeting at least 7 days prior to the meeting

   iii)

Stocks	Stock leader
Cod (Gadus morhua) in subdivisions 22–24, western Baltic stock (western Baltic Sea)	Marie Storr-Paulsen
Cod (Gadus morhua) in subdivisions 24–32, eastern Baltic stock (eastern Baltic Sea)	Margit Eero

The Benchmark Workshop will report by 8 March 2019 for the attention of ACOM.

Annex 4: Working Documents

WD 1 - Proxies of age–length-key for eastern Baltic cod to inform recent growth in stock assessment models
Margit Eero, DTU Aqua

WD 2 - Standardized Length Based Survey Indices for Eastern and Western Baltic Cod.
Casper W. Berg and Kasper Kristensen

WD 3 - Application of the egg production method to estimate stock trends and spawning-stock biomass of Eastern Baltic cod
Fritz Köster, Bastian Huwer, Gerd Kraus, Rabea Diekmann, Margit Eero, Andrei Makarchouk, Serra Orey, Jan Dierking, Piotr Margonski, Jens Peter Herrmann, Jonna Tomkiewicz, Daniel Oesterwind, Paul Kotterba, Rüdiger Voss

WD 4 - Baltic cod larvae index - Description of sampling and index calculation method
Bastian Huwer and Fritz Köster

WD 5 - Historical growth estimates of Eastern Baltic cod from tagging data.
Monica Mion, Alessandro Orio, Roman Motyka, Annelie Hilvarsson, Krzysztof Radtke, Karin Hüssy, Maria Krüger-Johnsen, Kate McQueen, Uwe Krumme, Maris Plikshs and Michele Casini.

WD 6. Estimation of natural mortality for Eastern Baltic cod
Massimiliano Cardinale

WD 7. Fits to age, length and conditional age at length (ALK) data for final Stock Synthesis run.

WD8 - Workshop on use of recreational fisheries catch data in stock assessment of WB cod

WD 9 - Workshop on separating Eastern and Western Baltic cod for stock assessment

WD 10 - Stock splitting of western and eastern Baltic cod in SD 24
Franziska Schade, Peggy Weist, Uwe Krumme

WD 11 - Note on a method for estimating stock mixing proportions by otolith shape
Christoffer Moesgaard Albertsen

WD 12 - Issue list Western Baltic Cod

WD 13 - Western Baltic cod catch data
Margit Eero, Marie Storr-Paulsen

WD 14 - German recreational catch estimation
Harry V. Strehlow and Simon Weltersbach

WD 15 - Documentation of data collection and analysis

WD 16 - Reconstructing Swedish cod catch numbers-at-age for tour-boat and private data in SD23
Andreas Sundelöf, Hege Sande, Esha Mohamed

WD 17 - Maturity of Western Baltic cod

WD 18 - Independent western Baltic cod recruitment index: juvenile cod data from commercial poundnets
Kate McQueen, Uwe Krumme

WD1- Proxies of age-length-key for eastern Baltic cod to inform recent growth in stock assessment models

Margit Eero, DTU Aqua

Summary

Three possible approaches to proxy Age-Length-Key (ALK) were developed that could be used to inform recent growth in stock assessment models. An ALK is needed in Stock Synthesis model to inform growth parameters, which are subsequently used to convert length distributions to age. The following options of ALK data were considered to be used in SS model:

ALK Scenario	Years	Basis	Sample size info
ALK1, Q4& Q1	1998-2006	Annual BITS age readings	Numbers of age readings in a given year
	2007-2009	Mean of 2002-06 & 2010-14	The total number of age readings in 2002-2006 divided by 5, to account for pooled data over 5 years.
	2010-2014(2017)	Constructed based in drivers/indicators	Same values as for 2007-2009 were used
ALK2, Q4	1998-2017:	Annual age readings of SWE/POL/LAT (BITS + Commercial)	Numbers of age readings in a given year
ALK3, Q4 & Q1	1998-2017	Annual age readings of all countries (BITS)	Numbers of age readings in a given year

The mean length at age derived from the ALKs 1-3 are shown in Fig 1 and 2 for Q4 and Q1 data, respectively.

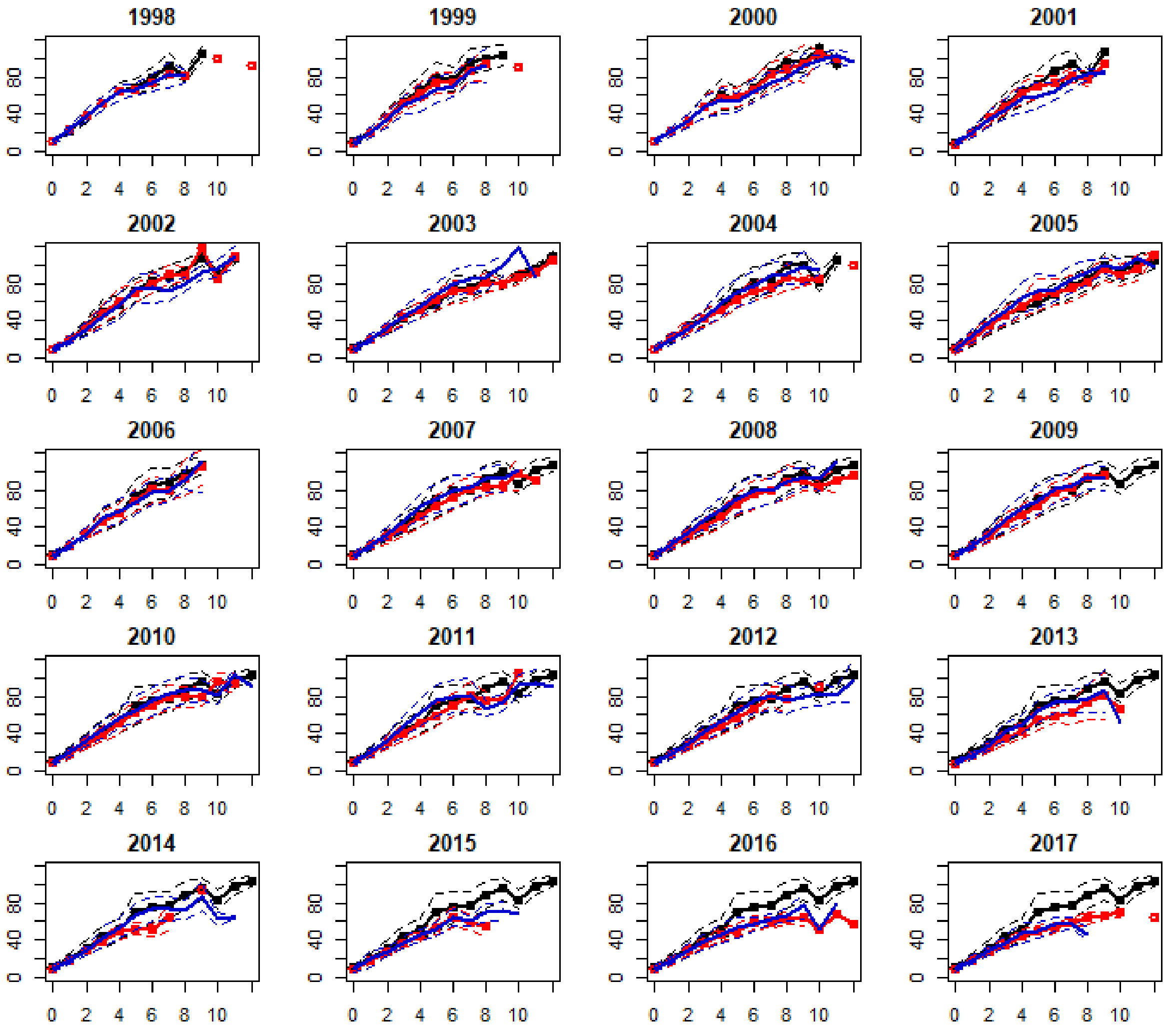


Fig. 1. **Black lines**: constructed ALK from 2006 onwards (ALK1). **Blue lines:** Mean length at age based on average ALK of SWE, POL, LAT (BITS Q4 data plus commercial data for >50cm cod) (ALK2); Red lines: Mean length at age (LAA) based on average ALK of all countries (BITS Q4 data) (ALK3). All based on Q4 data (individual sample data only, not raised to the population).

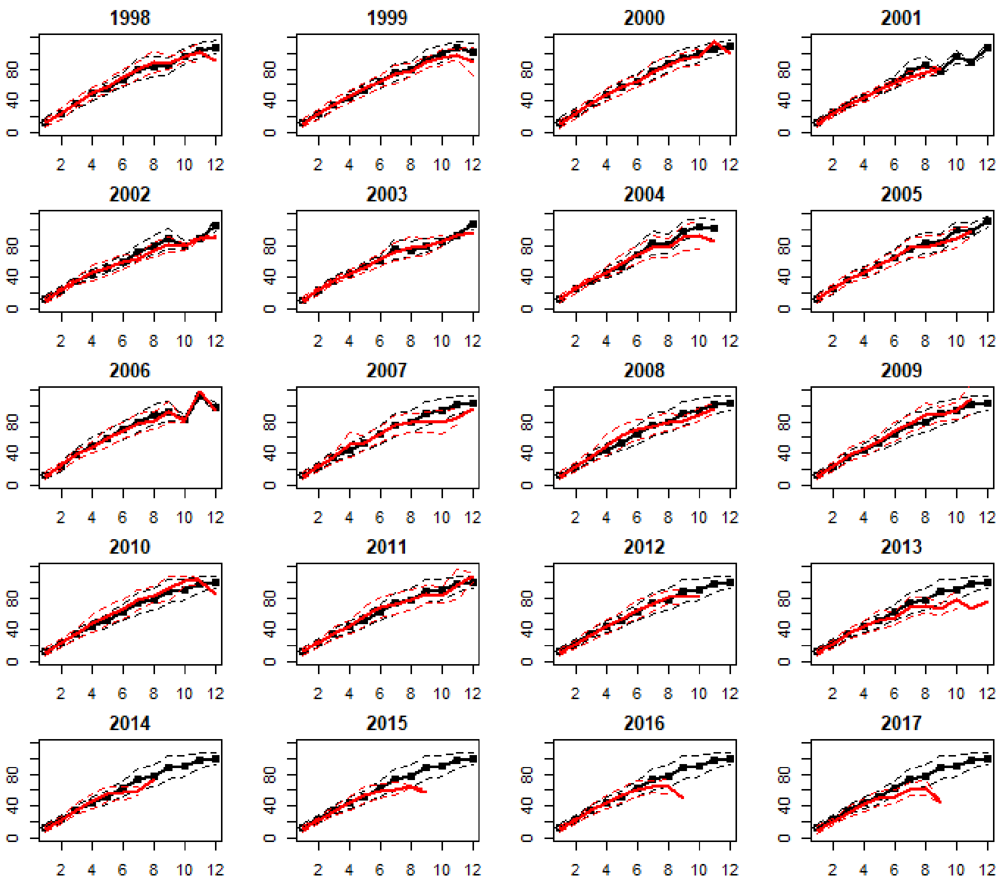


Fig. 2. **Black lines**: constructed ALK from 2006 onwards (ALK1); Red lines: Mean length at age (LAA) based on average ALK of all countries (BITS Q1 data) (ALK3). All based on Q1 data (individual sample data only, not raised to the population).

Background

It is recognized that age reading of EBC has always been difficult, introducing uncertainty in stock assessment (Hüssy et al. 2016). However, aging errors were not found to significantly influence the main conclusions in terms of stock status before the 2000s (Reeves 2003). Also, the diagnostics of stock assessment models did not indicate major inconsistencies in input data until the late 2000s, when retrospective bias in the assessment increased and eventually resulted in the failure of an age based assessment in 2014 (ICES 2014). As precision of age estimation is known to decrease during unfavourable environmental conditions which affect fish growth (Yaragina et al., 2011), the increase in ageing problems may be the result of changing growth conditions in later years (Eero et al., 2015). The age data from traditional age readings for before 2007 are considered still applicable for stock assessment purposes (ICES WKIDEBCA 2018). The ALKs described in this doc focus on providing proxies for informing growth from 2006 onwards.

ALK 1: based on observed change in drivers/indicators for growth

1.1 Mean length at age in 1991-2006

Age and length information for individual cod from BITS surveys were used to obtain annual ALKs for years 1991-2006. Smoothed ALKs were fitted, using *fitALK* function in DATRAS R package. The reason for using smoothed ALKs is to better detect differences in growth between time periods, as raw data is often associated with a lot of noise, especially for older ages. For younger ages, smoothed and raw ALK are similar (Fig. 1.1).

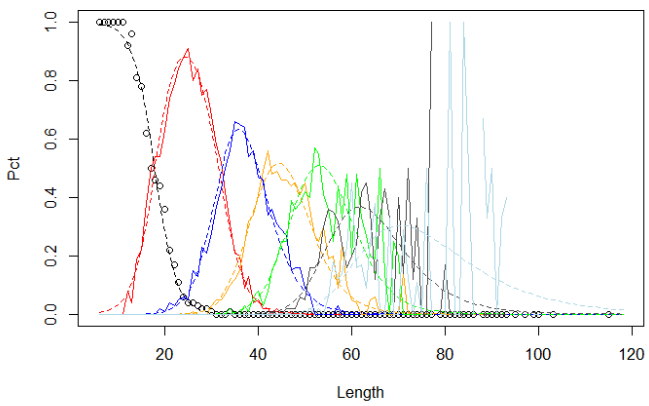


Fig. 1.1 Example of raw (solid line) and smoothed (stippled line) ALK (based on Q1 2005 data).

1.2 A proxy for the change in mean length at age since 2006

Drivers & indicators for change in growth

A number of changes in the Baltic ecosystem and in the cod stock have taken place in last decades, including a decline in nutritional condition of cod, reduced size at maturation and intensified hypoxia (Fig. 1.2). These changes are hypothesized to be associated with reduced growth (summarized in ICES WKBEBCA 2017), though direct relationships have not been demonstrated. Thus, there is a general consensus among the experts that the growth of EB cod must have declined as a result of the changes seen in cod biology and in the Baltic ecosystem (ICES WKIDEBCA 2018).

The relative importance of the potential drivers/indicators for cod growth (condition, size at maturation, hypoxia) are not known, however major trends in respective time series are relatively similar, showing a decline especially in the period from late 1990s to 2010s, with some levelling off afterwards. Each of these time series was standardized (by subtracting the mean and dividing by standard deviation), and the standardized time series were subsequently averaged, to obtain an overall index (Fig. 1.2). The magnitude of change in this index from 1994-1997 to 1998-2001 was similar compared to the magnitude of change in the period from 2002-2006 to 2010-2014. The time periods were chosen based on visual inspection of the average index representing changes in drives/indicators of cod growth, so that the values within a time period were as stable as possible. No information on the type of relationship between these potential drivers/indicators and cod growth is available, and if present, such relationship would possibly not be linear. However, at lack of any other information, a simple assumption was made that the change in cod growth in the period from the average in 2002-2006 to the average in 2010-2014 is similar to the observed growth change from the average level in 1994-1997 to the average in 1998-2001.

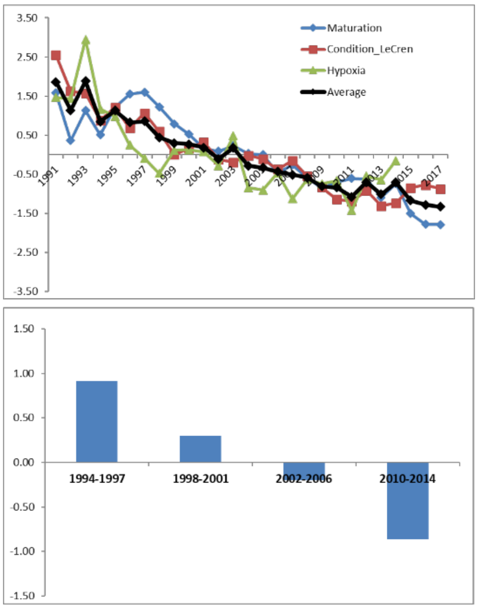


Fig. 1.2. Upper panel: Standardized time series of size at first maturation (L50) of cod, average condition of cod (Le Cren K) (estimated from BITS Q1 data in DATRAS) and extent of hypoxic areas (Casini et al. 2016). The time series of hypoxia is reversed to follow the same direction of trend as the other variables. The black line shows the average of the 3 standardized time series. Lower panel: The average index (shown as black line in the upper panel) averaged over the defined time periods.

Mean length at age by time periods: Quarter 1

The annual smoothed ALKs for Q1 were averaged over the years within the periods 1994-1997, 1998-2001, 2002-2006, corresponding to the time periods used for comparing changes in potential drivers/indicators for cod growth (Fig. 1.2). The resulting average length frequency distributions of age groups in the three time periods are shown in Fig. 1.3. As a next step, mean length at age was calculated for each age group, in the three time periods (Table 1).

The difference in mean length at age between time periods was calculated by dividing average mean length at age in 1998-2001 with the average mean length at age in 1994-1997, for each age group. Mean length at age of younger ages (1-2) was similar in the two time periods. The difference in mean length at age between the two periods was largest for middle age groups (3-5) (up to 10%), and less for older ages (6+) (2-6%). As an average for ages 3-9 (older fish not included due to few individuals), mean length at age was ca 6% lower in the later time period (Fig. 1.4). A similar proportional change in mean length at age was assumed in the period from 2002-2006 to 2010-2014.

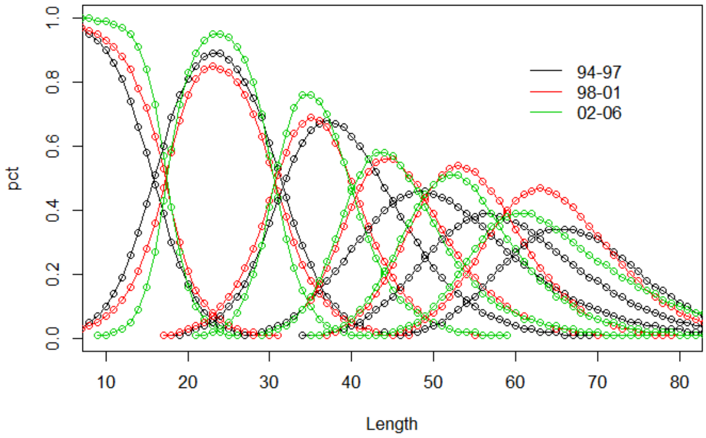


Fig. 1.3. Average smoothed length frequency distribution of cod age-groups (in percentage) in periods 1994-1997, 1998-2001 and 2002-2006, based on Q1 BITS data for SD 25-28. The length distributions are shown for ages 1-6. Individual sample data only, not raised to the population.

Table 1. Mean length at age (cm) in different time periods, based on BITS Q1

Age	1994-1997	1998-2001	2002-2006
1	11.7	12.2	11.9
2	23.9	24.1	24.2
3	39.4	35.8	36.3
4	51.8	46.4	46.3
5	59.6	54.6	54.3
6	67.7	65.0	65.7
7	79.1	77.7	77.5
8	89.3	84.1	82.3
9	92.3	90.3	91.2
10	99.6	98.2	96.7

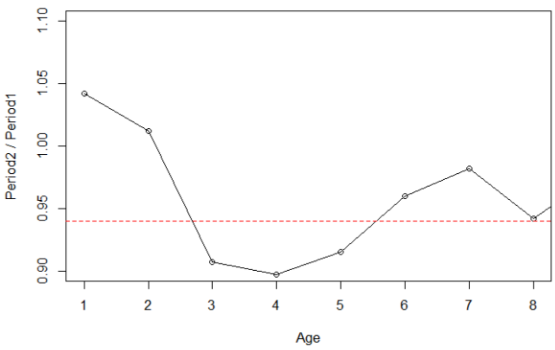
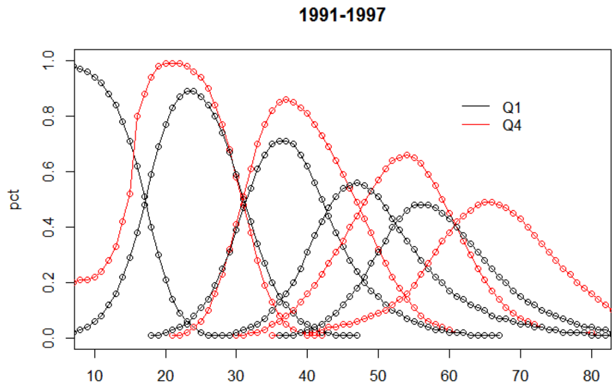
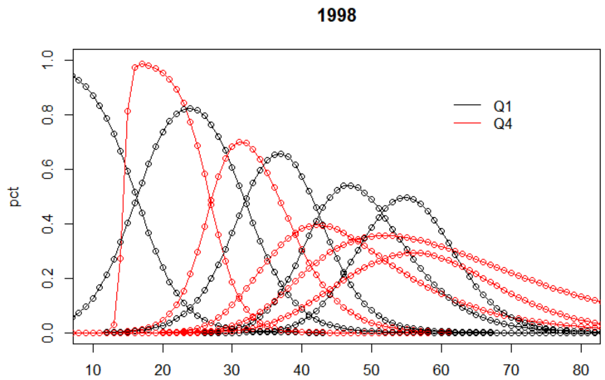


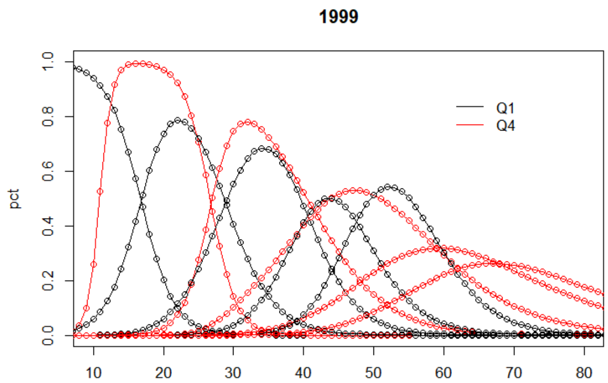
Fig. 1.4. Ratio between average mean length at age in 1998-2001 (Period2) and average mean length at age in 1994-1997 (Period1).

Mean length at age by time periods: Quarter 4

Similar analyses of changes in mean length at age over time as for Q1 were attempted for Q4. However, the comparison of ALK for Q1 and Q4 revealed that the age data for Q4 is ambiguous for years before late 1990s, i.e. the years when BITS Q4 data has also not been used for stock assessment, and the sampling sizes are relatively low. The length at age for Q4 in 1991-1997 (Fig. 1.5) showed very large values compared to Q1 data (e.g. peak at ca 55 cm for age 3 in Q4 compared to a peak at around 45 cm for age 4 in Q1 in 1991-1997). These inconsistency between quarters are less or not apparent in the data for the period 1998-2006. As the estimates for Q1 are based on a larger number of individual age readings, and show more realistic values, the Q4 age data for 1991-1997 were not further used. Thus, similar change in mean length at age from 2002-2006 to 2010-2014 as estimated for Q1, was assumed for Q4.







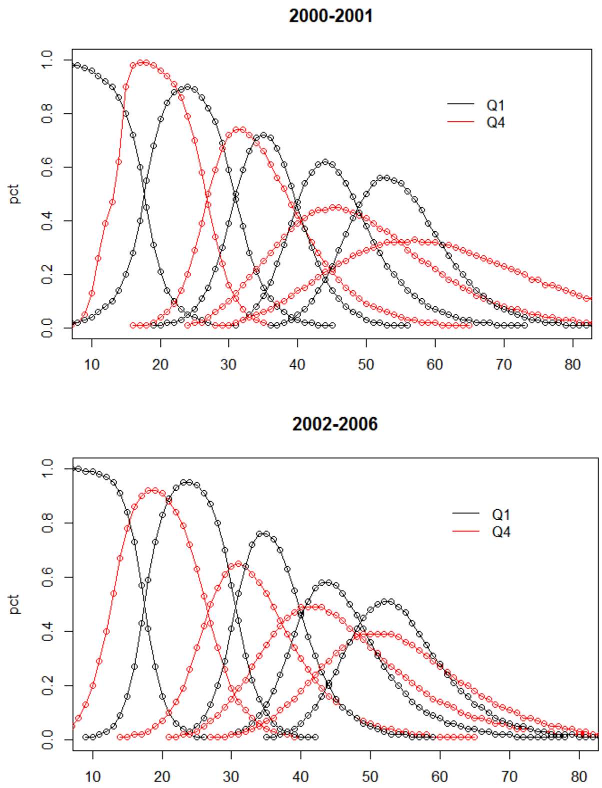


Fig. 1.5. Comparison on average smoothed length frequency distribution at age in Q1 and in Q4 in the periods 1991-1997, 1998, 1999, 2000-2001 and 2002-2006. Data are shown for ages 1-4 in Q4 and for ages 1-5 in Q1. This implies that e.g. length at age 2 in Q4 is expected to be larger than the length at age 2 in Q1, but smaller than the length at age 3 in Q1. Individual sample data only, not raised to the population.

1.3 Consistency of the proxy change in mean length at age with other growth indications

Given the unknown relationships between cod growth and the drivers/indicators for growth described above, alternative indications of the magnitude of growth change in later years were explored. Length frequency distribution (LFD) data comparing the progression of two relatively stronger year-classes formed in 2003 and 2011 (based on Köster et al. 2017) were used as an indication of growth change in this period. The presence of a stronger year-class following weaker ones makes it possible to follow a specific year-class as peaks in LFD data, at least for younger ages. Due to low representation of <20 cm cod in BITS survey catches, these relatively stronger year-classes from 2003 and 2011 are first visible in Q4 survey data as peaks in 2004 and 2012 (corresponding to age 1) and in Q1 survey data as peaks in LFD in 2005 and 2013, (corresponding to age 2)(Fig. 1.6). In subsequent years (2005 and 2013 for Q4 data and 2006 and 2014 for Q1 data), movement of these peaks towards larger sizes is detectable, corresponding to the particular year-classes at age 2 (Q4) or age 3 (Q1) (Fig. 1.6). The average length of cod corresponding to these peaks was calculated as the average of the 3 highest values in LFD:

	Length at age 1 (Q4)	Length at age 2 (Q1)	Length at age 2 (Q4)	Length at age 3 (Q1)

2003 YCL	21 cm	23 cm	31 cm	31 cm
2011 YCL	23 cm	24 cm	29 cm	29 cm

Length at smallest age (age 1 in Q4 and age 2 in Q1) does not appear to be lower for the 2011 year-class compared to 2003 year-class. In fact, mean length at age is somewhat higher for the later year-class for these youngest ages. However, this could be related to the long spawning time of the EB cod, as if the survivors originate from early in the spawning season, these are expected to have a larger length at age compared to those originating from the end of the spawning season. Mean length at age in subsequent year was however by 2 cm lower for the 2011 year-class (31 cm) compared to 2003 year-class (29 cm). This was consistent both for Q1 and Q4 data, implying a difference of 6 % (29/31=0.94) in mean length at age between in these periods. In terms of reduction in growth, this is a conservative estimate, as the 2011 year-class started out at a larger size than the 2003 year-class, as described above. The proportional change in mean length at age (6 %) indicated from length frequency data, is in line with the proxy on mean length at age derived based on change in drivers/indicators for cod growth.

The LFD of the same year-classes have also been used to validate estimation of growth of smaller cod based on otolith daily rings (Hüssy et al. 2018). The length distribution, following the 2003 year class, showing a peak at 23 cm at age 2 (in 2005) and at 31 cm at age 3 in the following year (Fig. 1.6), corresponds to an average growth of 8 cm. This corresponds to the growth estimated from the known-age samples' 2003 year class (8.3 cm) (Hüssy et al. 2018). For the 2011 year class, the peak in length distribution at age 2 at 24 cm and at 29 cm at age 3, corresponds to an annual growth of 5 cm. Thus, growth from age 2 to age 3 decreased from 8 to 5 cm in the 2011 year class. In known-age samples, growth between age 2 and 3 decreased significantly from 8.8 cm in the 1997 year class to 7.6 cm in the 2010 year class, however the growth for 2010 year-class was likely overestimated (Hüssy et al. 2018).

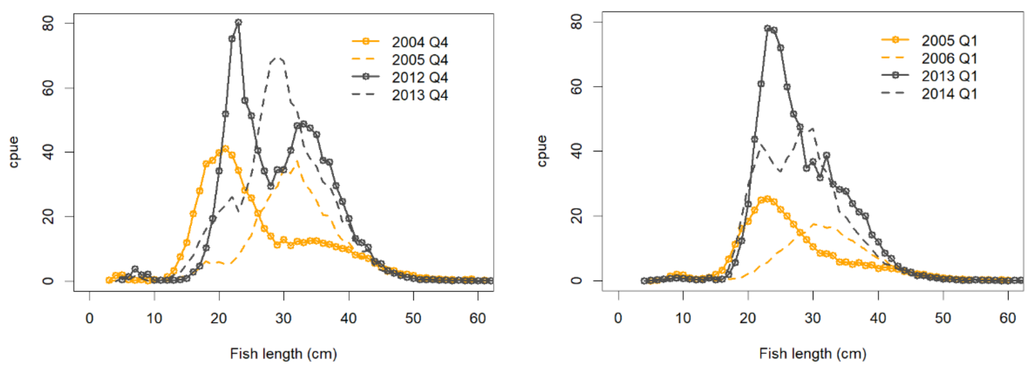


Fig. 1.6. Length frequency distribution of cod in SD 25-28 from BITS Q4 (left panel) and Q1 (right panel) survey in selected years.

1.4 Constructing proxy for Age-Length-Key for years 2006-2014

Based on the analyses described above, a 6% lower mean length at age in 2010-1014 compared to 2002-2006 was applied for all ages 3+ in Q1 and for ages 2+ in Q4. For younger age-groups (age 1-2 in Q1 and age 0-1 in Q4), a reduction in average length at age between the periods 1994-1997 and 1998-2001 was not obvious, based on age readings. Also, the length distribution data do not indicate lower mean length at age for these younger ages in later years, when comparing year-class from 2003 with the one from 2011. Thus, mean length at age of these ages (age 1-2 in Q1 and age 0-1 in Q4) in 2010-2014 was assumed to be the same as in 2002-2006.

The mean length at age for 2010-2014 was calculated by multiplying the mean length at age for 2002-2006 with 0.94, for ages 3+ (Q1) or 2+ (Q4). Subsequently, the corresponding difference in mean length at age in cm-s between 2002-2006 and 2010-2014 was calculated, shown in the table below, for Q1:

Age	1	2	3	4	5	6	7	8	9	10	11	12
diff_cm	0	0	2.2	2.8	3.3	3.9	4.7	4.9	5.5	5.8	6.2	6.3

Next, to obtain ALK for 2010-2014, the percentage length distribution within each age group in ALK for 2002-2006 was moved towards smaller size-classes by the cm-s shown in the table above (rounded to the closest cm). An example of this is provided below:

Length	a3 (2002-2006)	a3 (2010-2015)
19		0.01
20		0.01
21	0.01	0.02
22	0.01	0.03
23	0.02	0.05
24	0.03	0.08
25	0.05	0.13
26	0.08	0.2
27	0.13	0.29
28	0.2	0.41
29	0.29	0.54
30	0.41	0.64
31	0.54	0.72
32	0.64	0.76

This approach assumes that the range of variation in length distribution for a given age is the same as observed in 2002-2006, as the entire LFD for a given age is moved by certain cm-s. As the length distribution for younger ages in 2010-2014 were unchanged compared to the average for 2002-2006, while being adjusted for older ages, this caused some mismatch in ALK for 2010-2014 in a sense that the proportions of age groups for a given length did not sum to 1. This was subsequently corrected by multiplying the proportions of all age-groups for a given length by a common factor that made the age proportions for a given length to sum to 1. This resulted in average ALK for years 2010-2014 (Fig. 1.7 and 1.8). The average ALK for intermediate years (2007-2009) was calculated as an average of ALK in 2002-2006 and 2010-2014.

In subsequent years (2014 onwards), cod condition has remained stable, while size at maturation has further declined. Consequently, the size at age may have further declined from 2014 onwards.

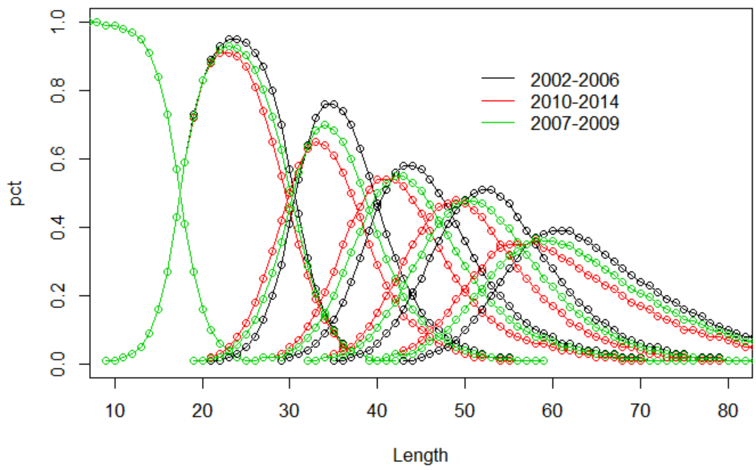


Fig. 1.7. Average relative smoothed length frequency distribution (in pct) of cod age-groups in 2002-2006, based on age reading data from BITS Q1 in SD 25-28, compared to the constructed ALK for 2010-2014 and 2007-2009. The length distributions are shown for ages 1-6. Individual sample data only, not raised to the population.

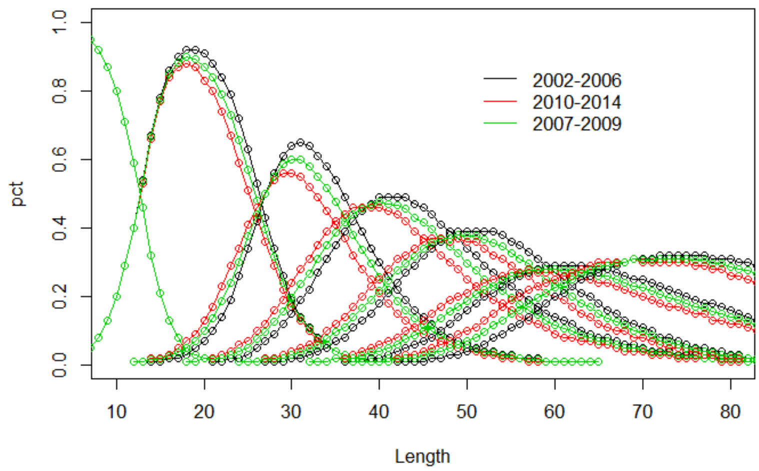


Fig. 1.8. Average relative smoothed length frequency distribution (in pct) of cod age-groups in 2002-2006, based on age reading data from BITS Q4 in SD 25-28, compared to the constructed ALK for 2010-2014 and 2007-2009. The length distributions are shown for ages 0-6. Individual sample data only, not raised to the population.

1.5 Consistency of CPUE at age when applying ALK1

CPUE at length was obtained from DATRAS database. For years before 2007, the annual ALK from otolith age readings of all countries was applied on CPUE at length to obtain CPUE at age values, thus a normal procedure that has regularly been applied in age based stock assessment. For years 2007-2017, the constructed average ALKs for periods 2007-2009 and 2010-2014 were applied on CPUE at length data in 2007-2009 and 2010-2017, respectively. This was done both for Q1 and Q4.

The internal consistency of the obtained time series was investigated by plotting the time series of indices at age *a* against the time series of indices for *a*+1 (Fig. 1.9). The internal consistency of the CPUE indices appears reasonable, and similar to the consistency of the times series that have been used in age based stock assessment by ICES previously.

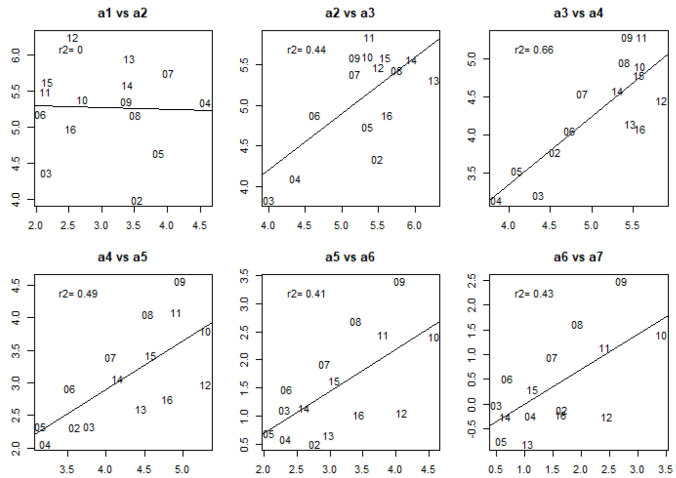


Fig. 1.9a Internal consistency of BITS Q1 survey, plotting indices at age *a* against the indices at age *a*+1. The numbers on the plot mark the year of capture of the younger age plotted on a particular panel. The age information is from ALK1 (age readings of all countries until 2006 & constructed ALK for later years; smoothed ALKs).

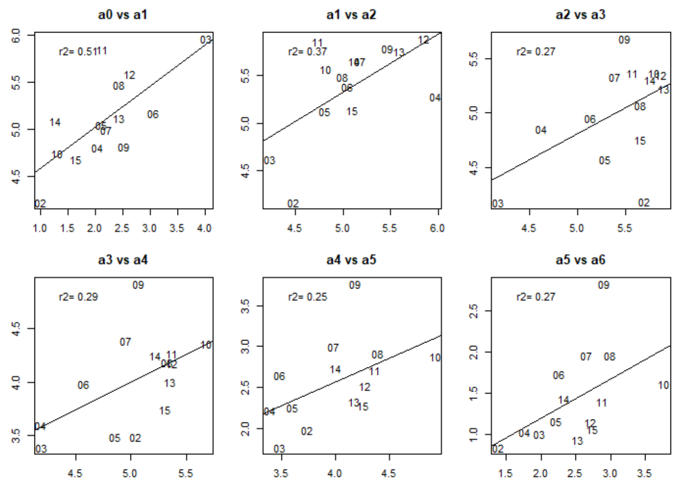


Fig. 1.9b. Internal consistency of BITS Q4 survey, plotting indices at age *a* against the indices at age *a*+1. The numbers on the plot mark the year of capture of the younger age plotted on a particular panel. The age information is from ALK1 (age readings of all countries until 2006 & constructed ALK for later years).

ALK2: based on national age readings of selected countries, BITS and commercial data combined

2.1 Selection of BITS data

The constructed ALK described as ALK1 was compared with data from national age readings with the aim to identify whether ALKs from some countries in later years are in line with the presumed change in growth represented by ALK1. The analyses focused on data for BITS Q4, as Q4 is preferred as input to Stock Synthesis model to inform growth, due to better representation of the smallest individuals. Figure 2.1 shows that annual age readings for BITS Q4 since 2006 are available from SWE, LAT, POL and LTU, while DEN and RUS have missing years in the later part of the time series.

ALKs differ for individual countries, also in years before 2006 (Fig. 2.2), when the combined international age data have been found useful for stock assessment purposes. The national differences in ALKs can be both due to differences in age interpretation as well as spatial differences in growth, as the cod that are age read in different counties are partly caught in different areas in the Baltic Sea. Individual countries ALKs also provide somewhat different perception of the change in growth from 2002-2006 to 2010-2015 (Fig. 2.3).

For these reasons, it was considered that if age readings are used as proxy to inform growth in stock assessment model in later years, only these countries' data should be used where information is available for all years. This is to avoid sudden changes in estimated growth resulting from changes in which countries' data are available for a particular year. This criteria is fulfilled for SWE, LAT, POL and LTU data. However, LTU ALK especially for latest period (2010-2015) shows unusual patterns with narrow even length intervals for all age groups, which has not been seen in other countries' data or in the time period before 2007 when the age information is still considered useful for stock assessment. For this reason, further analyses of national ALKs and their potential use to provide a proxy for informing growth in stock assessment focused on SWE, LAT and POL data.

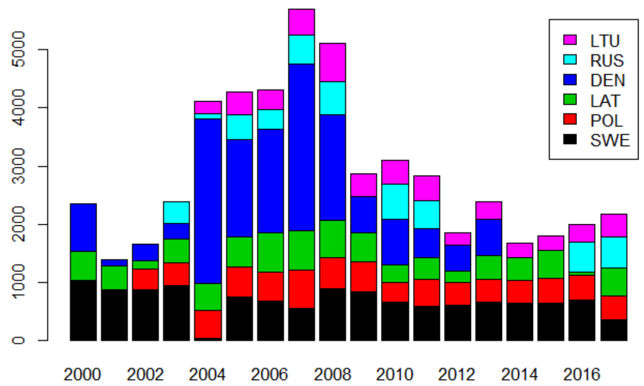


Fig 2.1. Number of otoliths age read by individual countries in Q4 BITS data, by year.

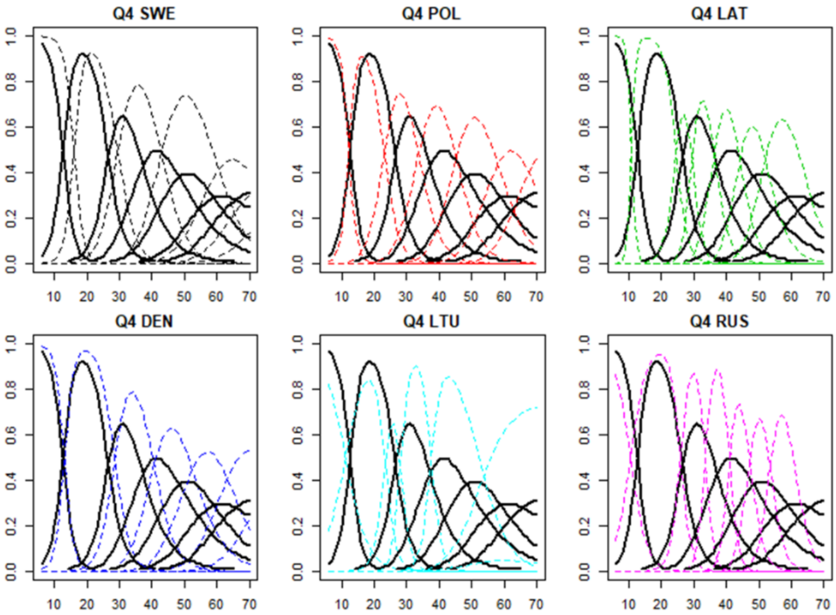


Fig. 2.2. Average ALK of individual countries in 2002-2006 (colored dotted lines) compared to the average of all countries in the same time period (solid black lines). Individual sample data only, not raised to the population.

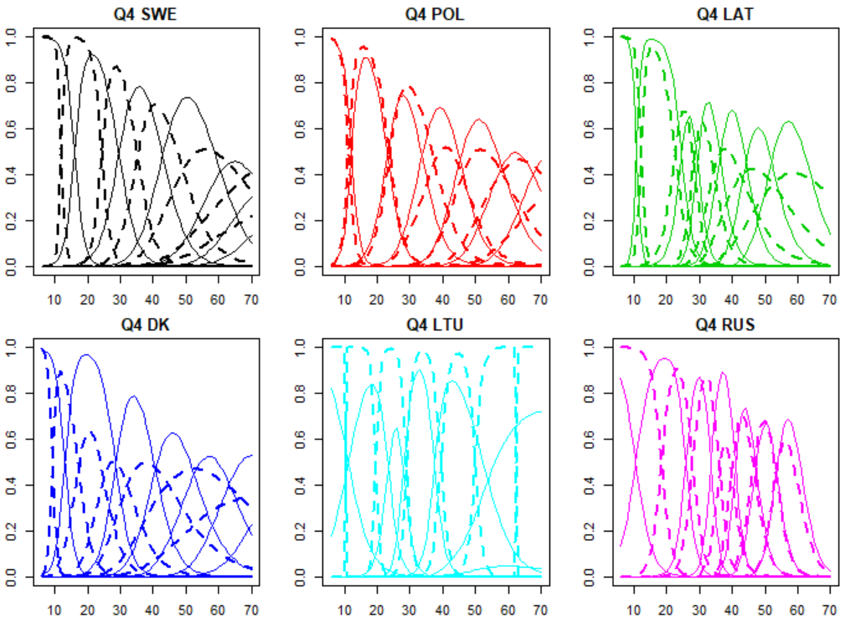


Fig. 2.3. Average relative length frequency distribution (in pct) of cod age-groups in 2002-2006 (solid lines) compared to 2010-2015 (broken lines), in Q4 BITS data, by country. The data are shown for age-groups 0-6. Individual sample data only, not raised to the population.

2.2 Length at age based on BITS age readings of selected countries

The ALKs for SWE, POL and LAT also differ (Fig. 2.4), however a combined smoothed ALK for these countries, at least for the length range below 50 cm, which covers major part of the stock in later years, is roughly in line with the constructed ALK based on changes in drivers/indicators for growth (described as ALK1) (Fig. 2.4)

We also analysed the average length at age (LAA) based on ALK (Fig. 2.5). For years 1998-2006, the average LAA based on all countries age readings was similar to that using only SWE/ POL/ LAT data. For 2006 onwards, the SWE/ POL/ LAT data was generally in line with the LAA from the constructed ALK (described in ALK1). However, the SWE/ POL/ LAT data suggest somewhat lower LAA for later years than the constructed ALK. However, as the number of larger/older cod in stock is very low in later years, age readings are basically not available for >60cm cod.

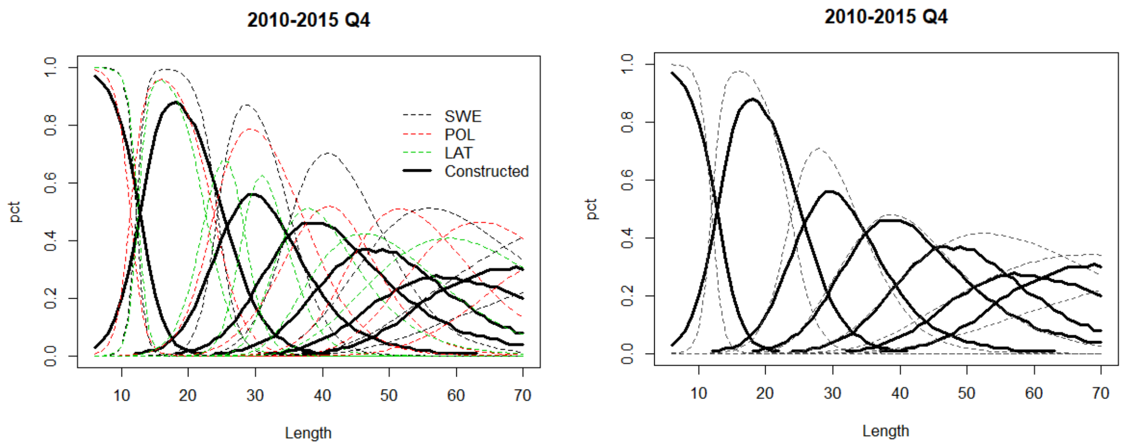


Fig. 2.4 Left panel: ALK by country compared to the constructed ALK (described under Option 1 in tis document). Right panel: combined ALK based on SWE, POL, LAT data shown on left panel, compared the constructed ALK (described under Option 1 in tis document). Individual sample data only, not raised to the population.

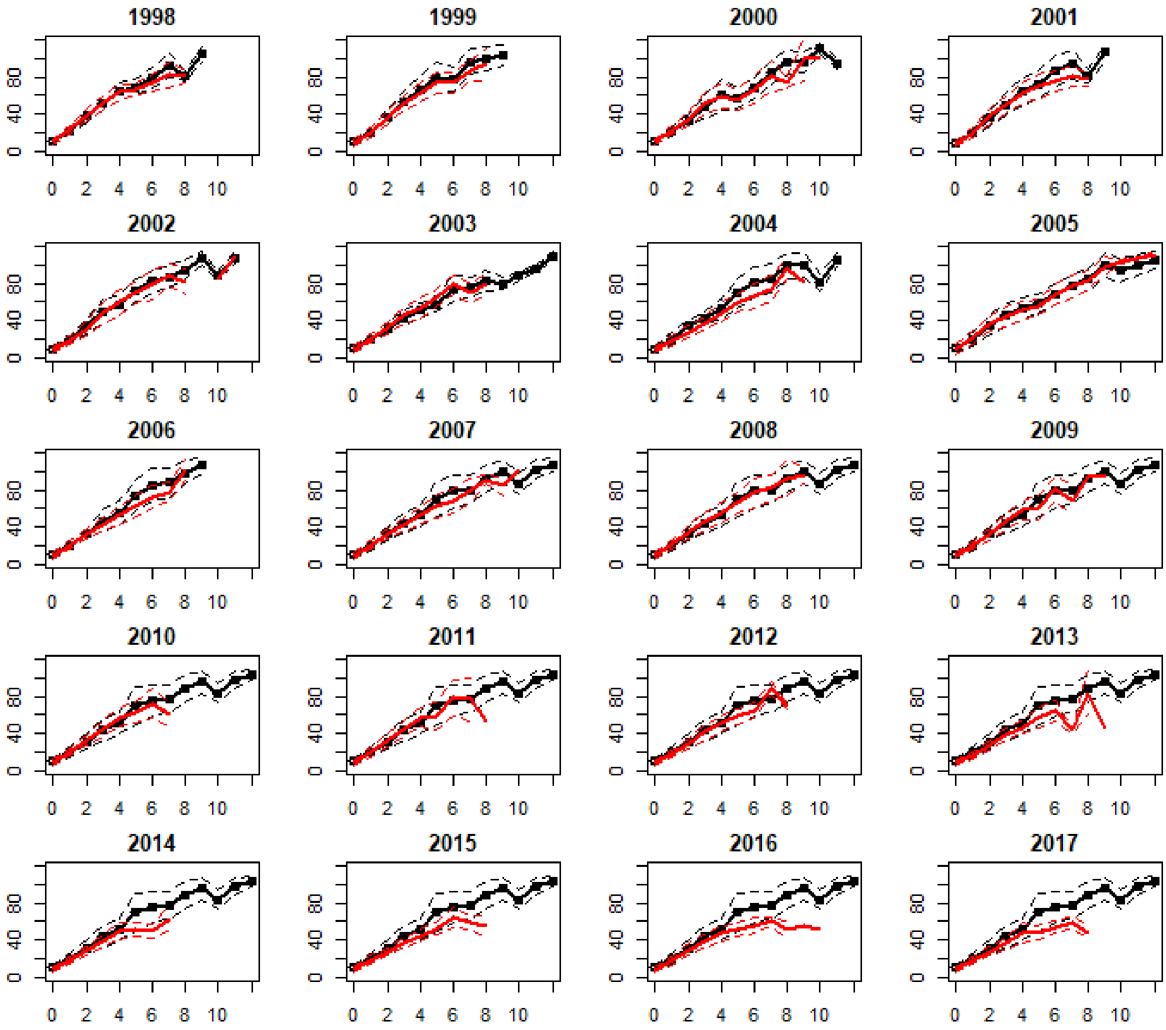


Fig 2.5. Mean length at age (LAA) based on average ALK of SWE, POL, LAT (BITS Q4 data) (red lines) compared to mean LAA based on all countries age readings (1998-2006) or the constructed ALK1 (2007 onwards) (black line). Individual sample data only, not raised to the population.

2.3 Incorporation of age readings from commercial data for selected countries

The number of larger cod in BITS ALK is very low (Fig. 2.6). Thus, to increase the abundance of larger individuals in ALK data, commercial data was additionally incorporated. Only cod >50 cm in length were included from commercial data. This is due to possible differences in ALK from surveys and commercial catches for smaller cod, related to gear selectivity.

The inclusion of commercial data for some years (e.g. 2010-2012) helped to get a better picture of LAA for older cod, that was generally in line with the constructed ALK (Fig. 2.7). However, for latest years (2013 onwards), few older cod are available also from commercial data (Fig. 2.6). The available age data suggest lower mean LAA for older cod for in latest years than estimates from the constructed ALK. However, the estimates for these older cod are based on very few data. Thus, there are no clear indications from these analyses whether growth of EB cod has been stable after 2012 or has declined further.

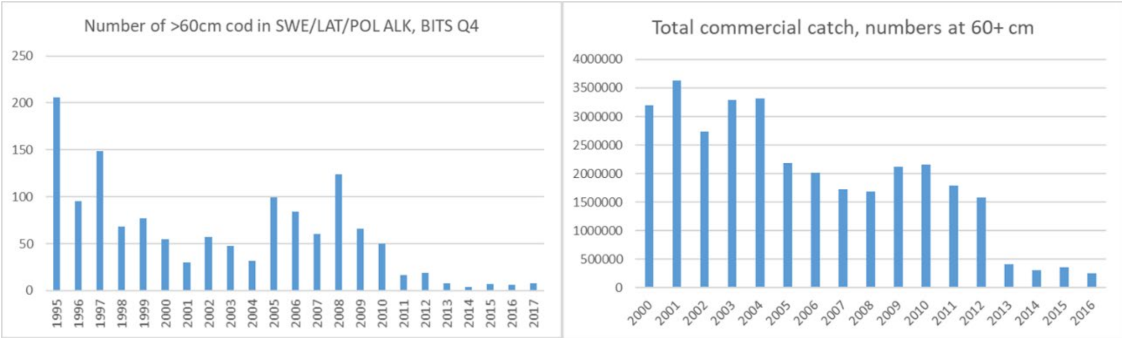


Fig. 2. 6. Left panel: Number of age readings in BITS Q4 survey (SWE/POL/LAT) for >60cm cod . Right panel: numbers of >60cm cod in commercial catch.

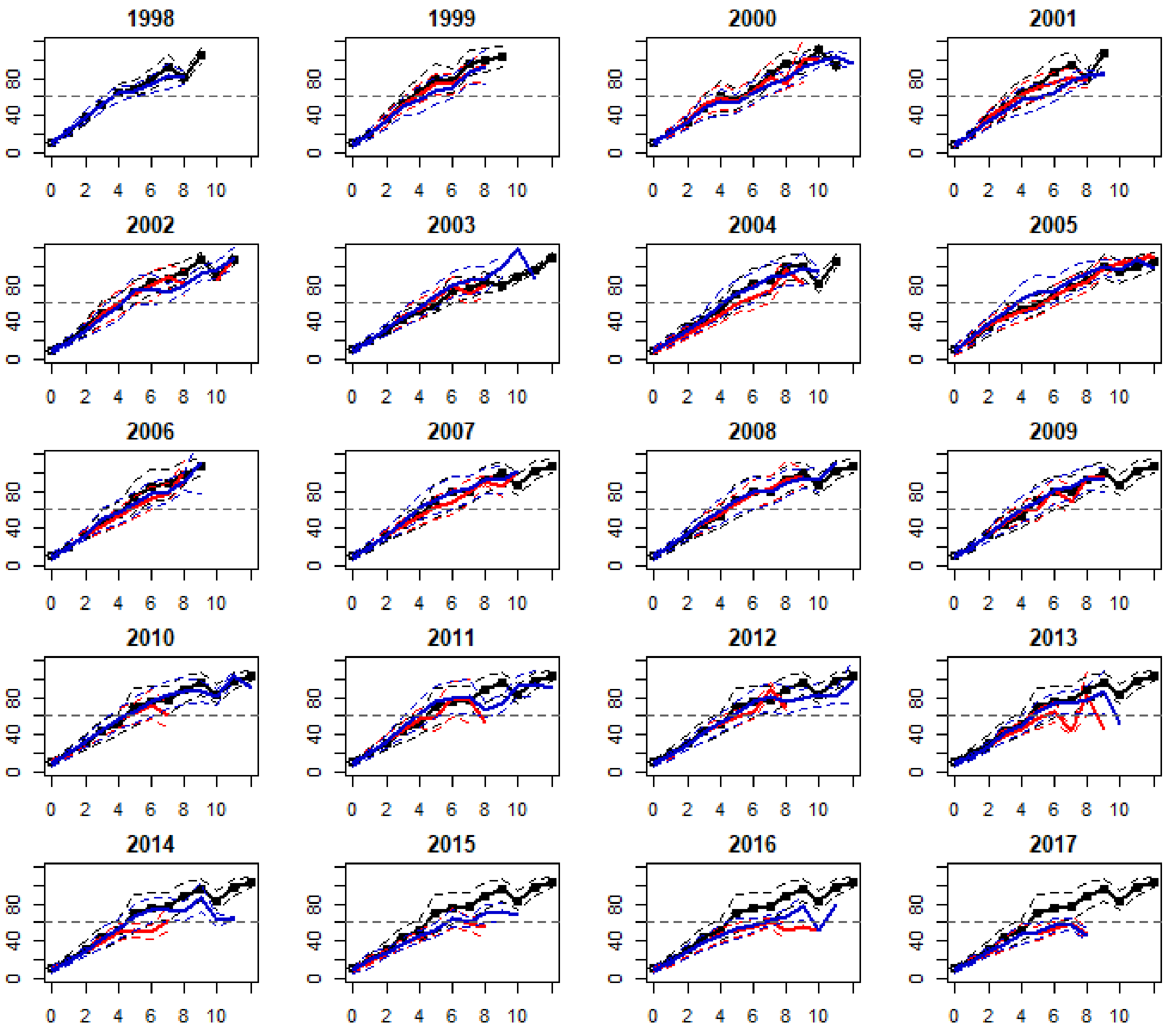


Fig. 2.7. Mean length at age (LAA) based on average ALK of SWE, POL, LAT (BITS Q4 data) (red lines); Mean length at age (LAA) based on average ALK of SWE, POL, LAT (BITS Q4 data plus commercial data for >50cm cod) (blue lines); Mean LAA based on all countries age readings (1998-2006) or the constructed ALK1 (2007 onwards) (black line). Individual sample data only, not raised to the population.

2.4 Consistency of CPUE-at-age when applying ALK2

To investigate he consistency of the proxy ALK2, this ALK was applied on BITS Q4 CPUE at length, and internal consistency of the resulting CPUE at age was analysed (Fig. 2.8). Two versions were investigating, i) using a separate ALK for each year, obtained as a combination of BITS and commercial data for SWE/POL/LAT; ii) keeping ALK constant from 2012 onwards, set to the values of 2012. The consistency of ALK2 was generally more poor than for ALK1 (especially for older ages). There was no major differences in consistency between option i) and ii) in applying ALK2, appart from ages 3-4, where the consistency was better with annual values (option i).

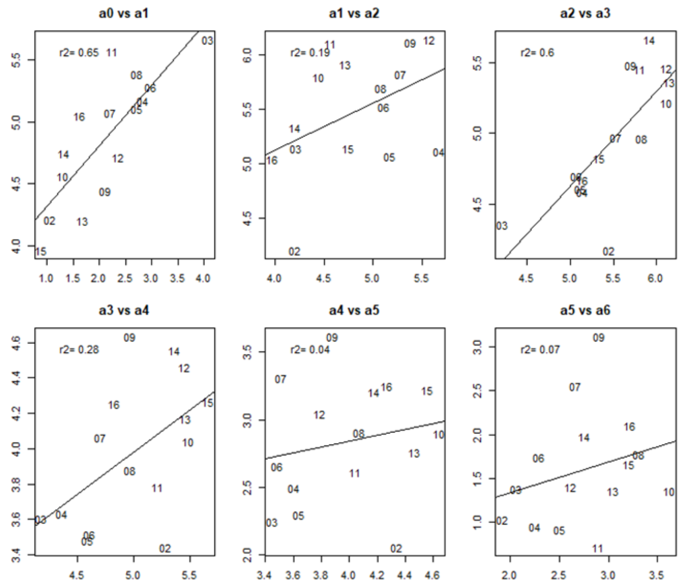


Fig. 2.8a. Internal consistency of BITS Q4 survey, plotting indices at age *a* against the indices at age *a*+1. The numbers on the plot mark the year of capture of the younger age plotted on a particular panel. The age information is from ALK2 (age readings from SWE/POL/LAT BITS and commercial data combined). Separate annual ALKs (non-smoothed) were applied for each year.

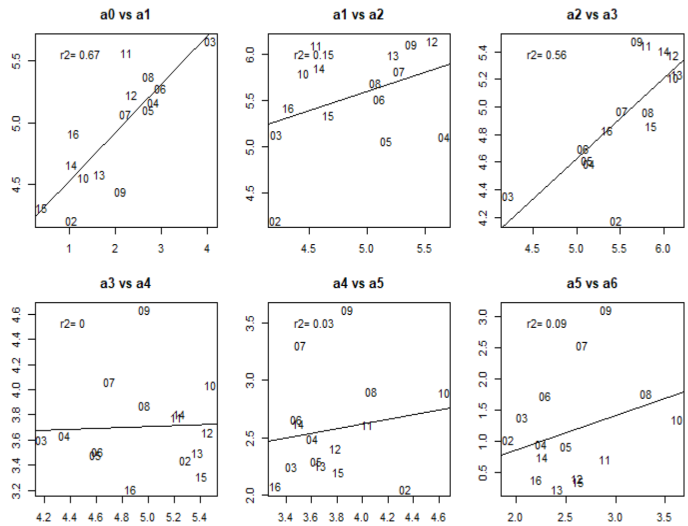


Fig. 2.8b. Internal consistency of BITS Q4 survey, plotting indices at age *a* against the indices at age *a*+1. The numbers on the plot mark the year of capture of the younger age plotted on a particular panel. The age information is from ALK2 (age readings from SWE/POL/LAT BITS and commercial data combined). Separate annual ALKs (non-smoothed) were applied for each year until 2012, and ALK for later years was set to be constant at the values of 2012.

ALK3: based on all available age readings from BITS

A third option for proxy ALK (ALK3) that was considered was to use all age information available in BITS. This option was explored because of the relatively poor performance of ALK2 in consistency analyses, especially for older ages (Fig. 2.8). The combined age readings for SWE/POL/LAT were mostly similar to ALK1 for younger ages, less so for older ones (Fig. 3.1). The ALK derived when combining all countries age readings indicated smaller fish at younger ages than ALK1, however for older ages the ALKs were more similar (Fig. 3.1).

3.1 Length at age based on BITS age readings of all countries

The analyses of mean length at age show similar picture as with ALK2. The ALK3 was similar to ALK1 until 2012, but indicates lower mean LAA for older cod in later years compared to ALK1. However, the ALK3 for older cod is still based on very few individuals. Thus, it remains unclear from these analyses whether growth of EB cod has been stable after 2012 or has declined further.

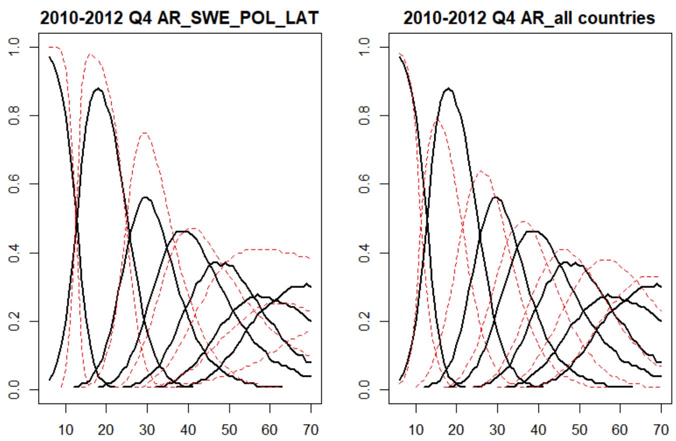


Fig. 3.1. Left panel: constructed ALK (ALK1, black line) compared to BITS age readings for SWE/POL/LAT (red dotted line), for 2010-2012 Q4 data. Right panel: constructed ALK (ALK1, black line) compared to BITS age readings for all countries (red dotted line), for 2010-2012 Q4 data. Individual sample data only, not raised to the population.

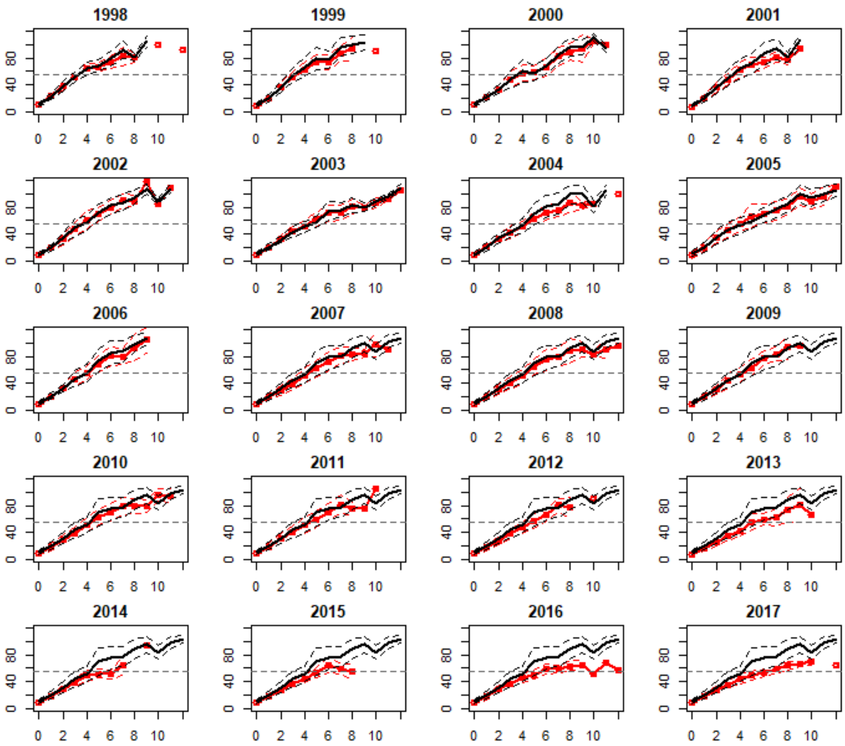


Fig 3.2. Mean length at age (LAA) based on average ALK of all countries (BITS Q4 data) (red lines) compared to the constructed ALK1 (2007 onwards) (black line). Individual sample data only, not raised to the population.

3.2 Consistency of CPUE-at-age when applying ALK3

To investigate he consistency of the proxy ALK3, this ALK was applied on BITS Q4 CPUE at length, and internal consistency of the resulting CPUE at age was analysed (Fig. 2.8). Two versions were investigating, i) using a separate ALK for each year; ii) keeping ALK constant from 2012 onwards, set to the values of 2012.

The consistency of ALK3 was generally better than for ALK2, especially for older cod, though was more poor for ages 1-2 (Fig. 3.3). There was no major differences in consistency between option i) and ii) in applying ALK3.

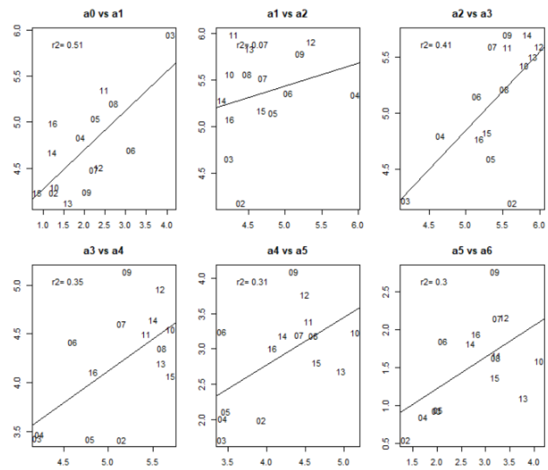


Fig. 3.3a. Internal consistency of BITS Q4 survey, plotting indices at age *a* against the indices at age *a*+1. The numbers on the plot mark the year of capture of the younger age plotted on a particular panel. The age information is from ALK3 (age readings from all countries from BITS). Separate annual (non-smoothed) ALKs were applied for each year.

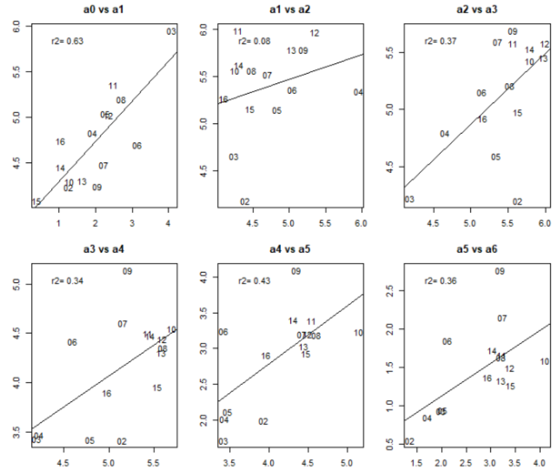


Fig. 3.3b. Internal consistency of BITS Q4 survey, plotting indices at age *a* against the indices at age *a*+1. The numbers on the plot mark the year of capture of the younger age plotted on a particular panel. The age information is from ALK3 (age readings from all countries from BITS). Separate annual ALKs were applied for each year until 2012, and ALK for later years was set to be constant at the values of 2012.

References

Casini M., Käll, F., Hansson, M., Plikshs, M., Baranova, T., Karlsson, O., Lundström, K., Neuenfeldt, S., Gårdmark, A., Hjelm, J. (2016). Hypoxic areas, density-dependence and food limitation drive the body condition of a heavily exploited marine fish predator. Royal Society Open Science. 3: 160416.

Eero,M.,Hjelm,J.,Behrens,J.,Buchmann,K.,Cardinale,M.,Casini,M.,Gasyukov,P.,Holmgren,N.,Horbowy,J.,Hüssy,K.,Kirkegaard, E., Kornilovs, G., Krumme, U., Köster, F. W., Oeberst, R., Plikshs, M., Radtke, K., Raid, T., Schmidt, J., Tomczak, M. T., Vinther, M., Zimmermann, C., and Storr-Paulsen, M. Eastern Baltic cod in distress: biological changes and challenges for stock assessment. – ICES Journal of Marine Science,72: 2180–2186

Hüssy. K., Eero, M., Radtke, K. 2018. Faster or slower: has growth of eastern Baltic cod changed? Marine Biology Research, DOI: 10.1080/17451000.2018.1502446

ICES 2017. Report of the Workshop on Biological Input to Eastern Baltic Cod Assessment (WKBEBCA). ICES CM 2017/SSGEPD:19

ICES 2018. Workshop on evaluation of Input Data to Eastern Baltic Cod stock Assessment (WKIDEBCA). 23-25 January 2018, Copenhagen, Denmark

Köster, F. W., Huwer, B., Hinrichsen, H. H., Neumann, V., Makarchouk, A., Eero, M., Dewitz, B. V., et al. 2017. Eastern Baltic cod recruitment revisited—dynamics and impacting factors. ICES Journal of Marine Science, 74: 3–19.

Reeves, S. A. 2003. A simulation study of the implications of age-reading errors for stock assessment and management advice. – ICES Journal of Marine Science, 60: 314–328.

Standardized Length Based Survey Indices for Eastern and Western Baltic Cod.

Casper W. Berg & Kasper Kristensen

November 20, 2018

1 Summary

2 Data

The data is coming from the BITS survey in the DATRAS database. The data set used to fit the model used for the standardization consists of almost the entire BITS database, although the indices of interest are computed for smaller sub-regions. Hauls with "HaulVal" codes "V", "A","N" and "C" and with "StdSpecRecCode" equal to 1 or 3 are included in the analysis. All gear types are included, except those that have been used for less than 120 hauls (this criterion excludes also pelagic trawls, see figure 7). Data from ICES areas 20 and 21 are excluded, except those below 56.5° latitude (i.e. a northern edge of one extra ICES square going into Kattegat, to improve estimates near the edge of the domain boundary). A few north eastern hauls (east of 22° longitude and north of 58.15° latitude) have also been excluded, because including these would lead to inclusion of many unsampled grid cells with low abundance in the model, which would increase computation time without improving the indices.

At some stations with no oxygen (HaulVal code "N") no haul was performed (haul duration 0 or `NA`). In order to inform the model that the abundance are likely to be low at these stations, the haul duration was set to 10 min for these hauls, i.e. one third of the conventional duration of 30 minutes. This means that these zero observations are included in the model estimates of abundance, but they are not given as much weight as an actual performed haul.

Although most length data are registered using 1 cm resolution, the data are divided into coarser size bins at the tails of the size spectrum prior to the analysis in order to reduce the computational load. The following size groups (first one is 9 cm and below, second one is 10 and 11 cm, etc.) are used:

```
[0,10), [10,12), [12,14), [14,16), [16,18), [18,20), [20,21), [21,22), [22,23), [23,24)
[24,25), [25,26), [26,27), [27,28), [28,29), [29,30), [30,31), [31,32), [32,33), [33,34)
[34,35), [35,36), [36,37), [37,38), [38,39), [39,40), [40,41), [41,42), [42,44), [44,46)
[46,48), [48,50), [50,52), [52,54), [54,56), [56,58), [58,60), [60,200)
```

2.1 Bathymetri

The Baltic Sea Bathymetric Database (BSBD, [1],500 m resolution) is used to define the domain of interest, and the survey indices are calculated by summing up standardized catch rates over all the grid cells. To speed up the calculations the resolution of the final grid was reduced to 2500 m. See figure 8 for a plot of the final grid.

3 Model

The statistical model estimates the expected catch of each size group s at each point in space x and time t for all gear types. The model is known as a log-Gaussian Cox model, which describes large-scale abundance fields and local patchiness using correlated log-Gaussian variables, and if these were known, the catch numbers would be Poisson distributed [4]. Unlike the model in [4] we do not assume any population dynamics structure (this is the job of the assessment model that uses the results of this analysis). The objective of our model is to provide standardized estimates of the total catch at length, i.e. the result of a virtual survey where hauls were taken at the exact same time in every grid cell using the same reference gear. Another way of formulating this is that we want the model to filter out the effects of changes in the survey design on the indices, most importantly gear effects and effects of unequal sampling intensity over the domain of interest. We also want to filter out random noise, which is also important especially for very patchy species where a single haul with a large catch can determine the value of an unfiltered survey index estimate. Rather than estimating a 4-dimensional field in (size, space x, space y, time) as in [4] we choose to estimate independent models for each size group in order to reduce the computational burden. The model is implemented in Template Model Builder [3].

3.1 Process equations

For a given size group the field $\eta(x, t)$ is a Gaussian zero-mean stochastic process which is correlated in space, and time. We assume a multiplicative correlation structure, such that the auto-covariance function of η can be factorized into two terms a spatial correlation ρ_{space}, and a time correlation ρ_{time}:

$$\text{Cor}\left(\eta(x,t) \cdot \eta(x + \Delta x, t + \Delta t)\right) = \rho_{\text{space}}(x, x + \Delta x) \cdot \rho_{\text{time}}(\Delta t)$$

Time is discretized in two steps per year (again to reduce computational demands). The correlation function ρ_{time} is as assumed to be that of an stationary AR(2) process, i.e. damped exponential and/or sine functions. This formulation is able to capture seasonal recurring patterns in the spatial distribution of a given length group. For example, the next year's distribution in quarter 1 can have a higher correlation with last year's quarter 1 than last year's quarter 4, even though the latter is closer in time. This is an extension compared to [4], where the correlation was exponentially decreasing function of time.

The geographical area discretized into cells in a uniform grid, assuming that the field $\eta(s, \cdot, t)$ is constant within each grid cell. Thus, the auto-covariance function ρ_{space} is replaced by a covariance matrix Σ_{space}. The precision matrix Q, which is the inverse of the covariance matrix Σ_{space}, is specified as follows:

$$Q_{ij} = \begin{cases} -q & \text{if cell } i \text{ neighbors cell } j \\ q \cdot (m_i + \delta) & \text{if } i = j \\ 0 & \text{else.} \end{cases}$$

Here, m_i is the number of neighbors of grid cell i; typically 4 but less for boundary cells. q and δ are parameters which are estimated. This precision structure implies that a cell is conditionally independent of a distant region, when conditioning on the cell's neighbors. The parameters q and δ can be transformed into a variance parameter σ^2 and a decorrelation range parameter H as follows:

$$\sigma^2 = \frac{1}{M}\mathrm{tr}\left(Q^{-1}\right) \quad , \qquad H = \frac{h}{\log(1 + \frac{\delta}{2} + \sqrt{\delta + \delta^2/4})} \quad . \tag{1}$$

Here M is the number of grid cells and the trace $\mathrm{tr}\left(Q^{-1}\right)$ is the sum of diagonal elements of Q^{-1}, so that σ^2 is the spatially averaged variance of the field $\eta(s, \cdot, t)$. h is the grid cell size. If the domain had been a one dimensional line, then σ^2 would be the variance while H would be the decorrelation range, so $\Sigma_{ij} = \sigma^2 \exp(-|i-j| \cdot h/H)$. Similar for a 2D spatial field this implies that the correlation decreases with distance traveled through water.

3.2 Observation equations

If N_i is the number of individuals in at a given size group in the ith haul, then

$$N_i \mid \eta(x_i, t_i) + \eta_0(i) \sim \text{Poisson}\Big(\exp\big(\eta(x_i, t_i) + \mathrm{t}(i) + \mathrm{Gear}(i) + f_1(\mathrm{depth}_i, \mathrm{Quarter}_i) + \mathrm{TimeOfDay}(i) + \log(\mathrm{HaulDur}_i) + \eta_0(i)\big)\Big)$$

where $\eta(x_i, t_i)$ is the random space-time field, and $\mathrm{t}(i)$ and $\mathrm{Gear}(i)$ are categorical fixed effect parameters of the time-step (discretized into quarterly bins) and gear. The "TVL" gear is chosen as the reference gear, such that the "TVL" gear effect is set to zero and all other gears effects are relative to this gear type. The f_1 function is a second degree polynomial (re-scaled to have mean zero over when evaluated in the observed depths) with distinct parameters by quarter. The maximum of the second degree polynomial (assuming a negative coefficient for the quadratic term) can be interpreted as a "preferred depth" for a given length group. The $\mathrm{TimeOfDay}(i)$ term maps the time of day into one of three categorical levels { 1 = Mid-day (10-14), 2 = early/late day (7-10 and 14-17), 3 = night (the rest) }. The catch is assumed to be proportional to the haul duration, so this is included as an offset (linear term with assumed coefficient of 1). Finally, there is a haul specific independent random effect η_0, the so-called nugget effect. The resulting Lognormal-Poisson mixture is similar to a negative binomial distribution, due to the similarity of the Lognormal and the Gamma distribution and because the negative binomial distribution is the same as a Gamma-Poisson mixture.

For further details about the statistical methodology the reader is referred to [4].

3.3 Index uncertainties

This subsection describes how the uncertainty (coefficient of variation) for the total catch in numbers (summed over space and length groups for each of the time points) is computed. The resulting CVs can be used as data weightings in the assessment model.

The survey index is obtained by summing the expected values of the observation equation in every grid cell at the times of interest using fixed values for all nuisance parameters (Gear, time of day, haul duration). Since models are fit independently by length group, the error distribution around the survey index estimate can also only be estimated for one length group at a time, i.e. correlations between length groups cannot be estimated within the model. We compute the uncertainty for the log of the survey index for a given length group using the delta method (through the ADREPORT functionality in TMB). The log transformation is chosen, because the log index is typically better approximated by a normal distribution than the index without takin logs.

An approximate expression for the variance of the logarithm of the total abundance of all length groups (in numbers and for a given point in time), $\mathbb{V}(\log \sum I)$, in a given time-step can be found using the delta method, given that we have an estimate of the variance of the log of the individual length groups $\mathbb{V}(\log I_l)$:

$$\mathbb{V}\left(\log\left(\sum I\right)\right) = \frac{I}{\sum I}^T \Sigma_{\log I} \frac{I}{\sum I} \tag{2}$$

where I is the vector of survey indices by length, and $\Sigma_{\log I}$ is the covariance matrix for the log of that vector $\log I$ (which is a diagonal matrix due to the independent model fitting). The square root of this expression (standard deviation of the log indices) is approximately equal to the CV of the total abundance, which may be used for data weighting in the assessment model.

4 Results

The survey indices obtained from the LGCP model are in overall agreement with the ICES indices in terms of the overall abundance trends (Figures 1 to 5). The largest discrepancies (and uncertainties) are found in the Q4 indices before year 2000 (see figure 3), because of reduced spatial coverage and mixture of gear types in this period. Two examples where the discrepancy is large are in 1994 Q4 and in 2008 Q4. At those time points ICES biomass estimates are 2.71 and 1.65 respectively (relative to a mean of 1), whereas the LGCP estimates are 0.87 and 2.45 respectively. The spatial plots (see Appendix) reveals that this discrepancy can be explained by the spatial coverage: in 1994 Q4 the majority of the assessment area was not covered, while in 2008 Q4 it seems that central areas predicted to contain high abundances by the LGCP model were sampled less intensively compared to other years.

ICES area 24 (see figure 9) is known to contain a substantial proportion of EB cod. Therefore figure 6 compares biomass indices (scaled to mean 1) computed from the LGCP model over two different regions, namely the standard EBcod assessment area (25-32) and one which also includes the eastern part of area 24 (area east of 13° longitude). It is seen that the indices are very similar, although the ratio between the two indices is above 1 around the time when the abundance of larger cod dropped drastically (year 2010), while it tends to be below 1 in the latest years. This implies that the decline in abundance is stronger in areas east of are 24. This could indicate that a small part of the drop in the standard indices can be explained by an increasing proportion of EB cod occupying area 24. The spatial plots confirm this pattern (e.g. Figure 23). The high concentration of abundances around the eastern and western stock border between area 24 and 25 is likely to cause problems due to mixing of the two stocks.

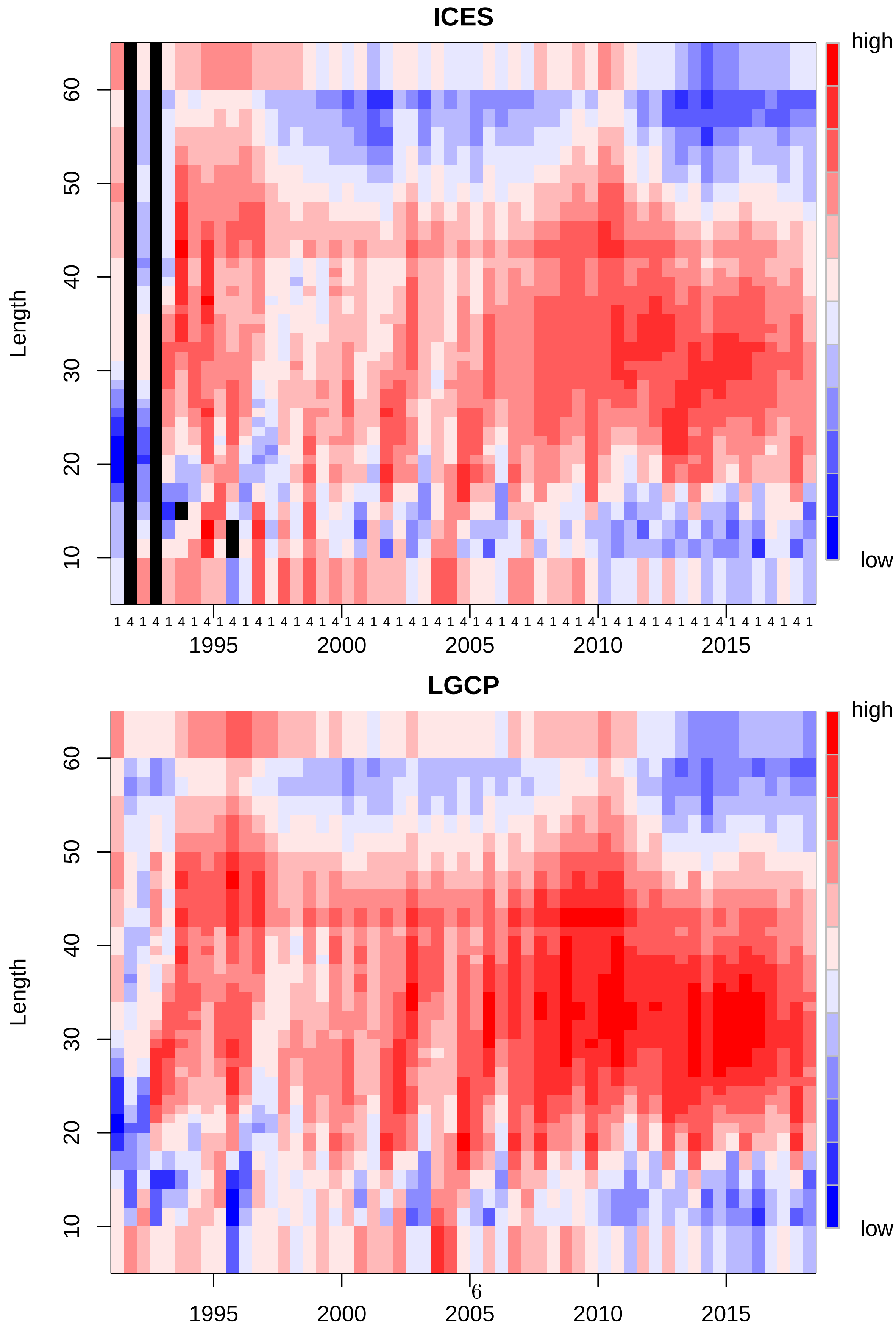


Figure 1: Top: ICES indices (aggregated to the same size bins as the LGCP index). Bottom: Estimated indices from the LGCP model. The color scale is equidistant on a logarithmic axis, and the x-axis alternates between quarter 1 and quarter 4. Black means missing (or zero).

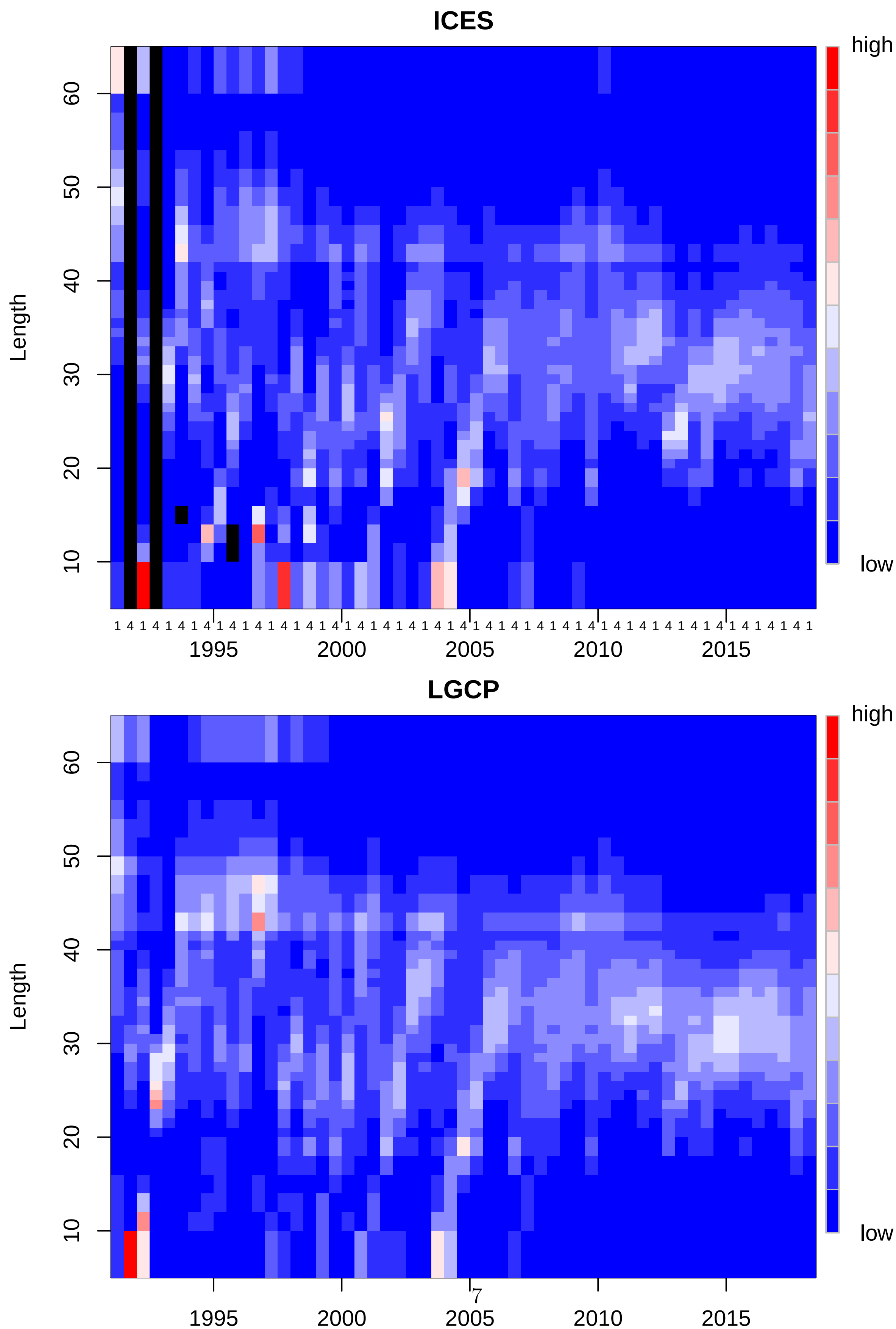


Figure 2: As figure 1, but rather than the log-abundances this figure shows the length proportions within years (i.e. each year has been divided by the sum on natural scale).

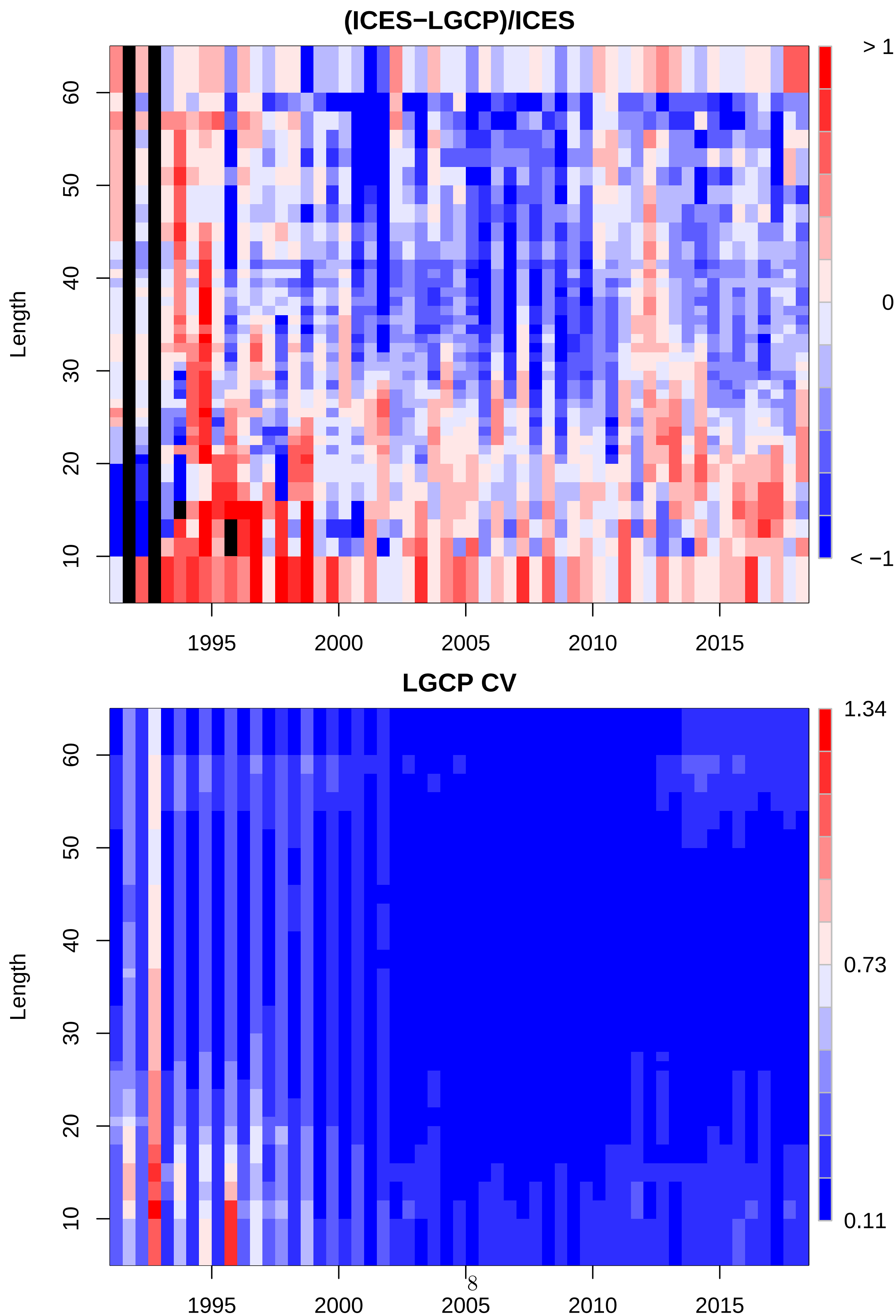


Figure 3: Top: Relative discrepancy between ICES and LGCP indices.
Bottom: Standard deviation of LGCP index on log scale, which is approximately equal to the coefficient of variation.

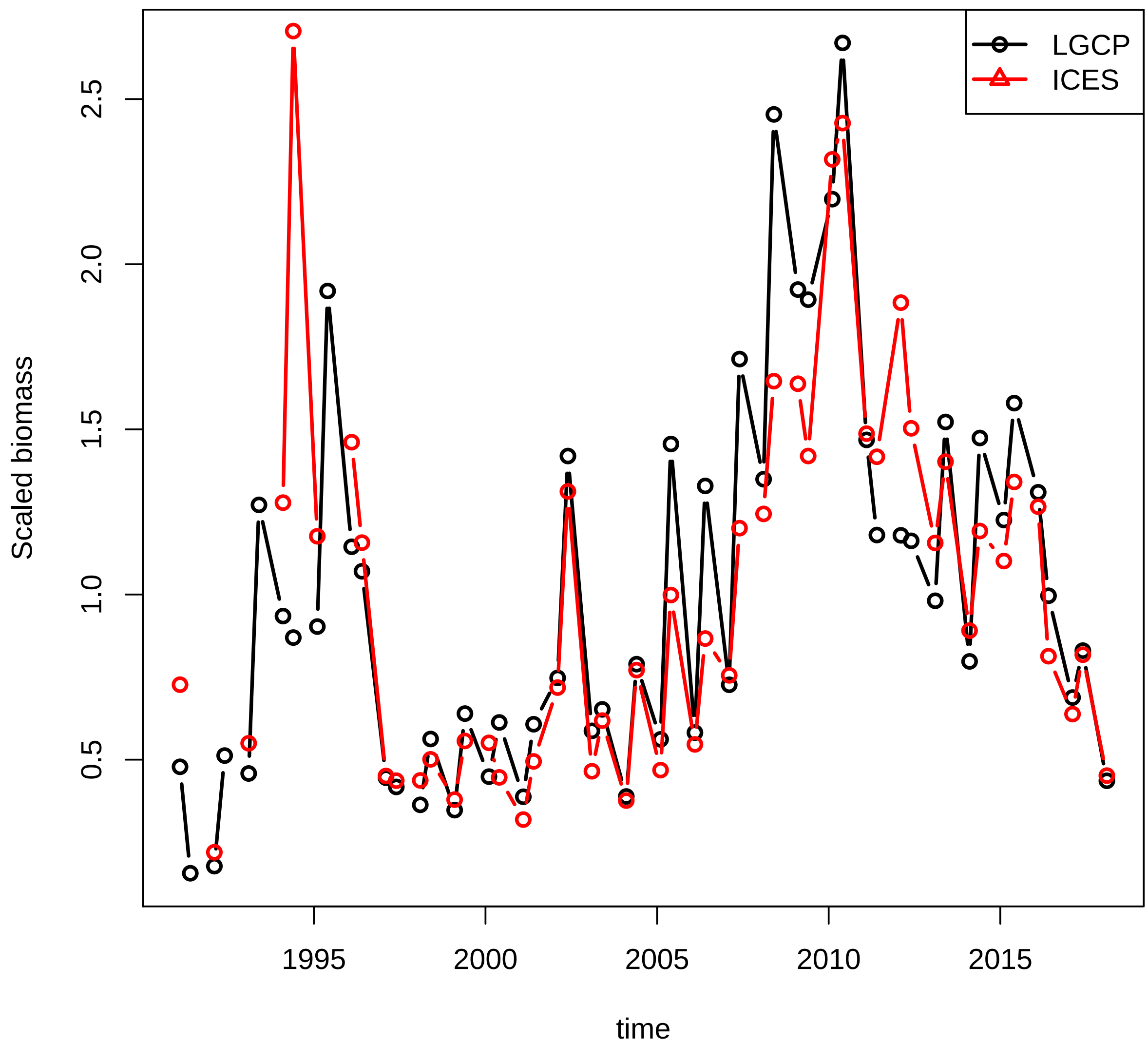


Figure 4: Time-series of total biomass (scaled to have mean 1 over time). Biomasses are calculated assuming a time-invariant length-weight relationship ($W = aL^3$).

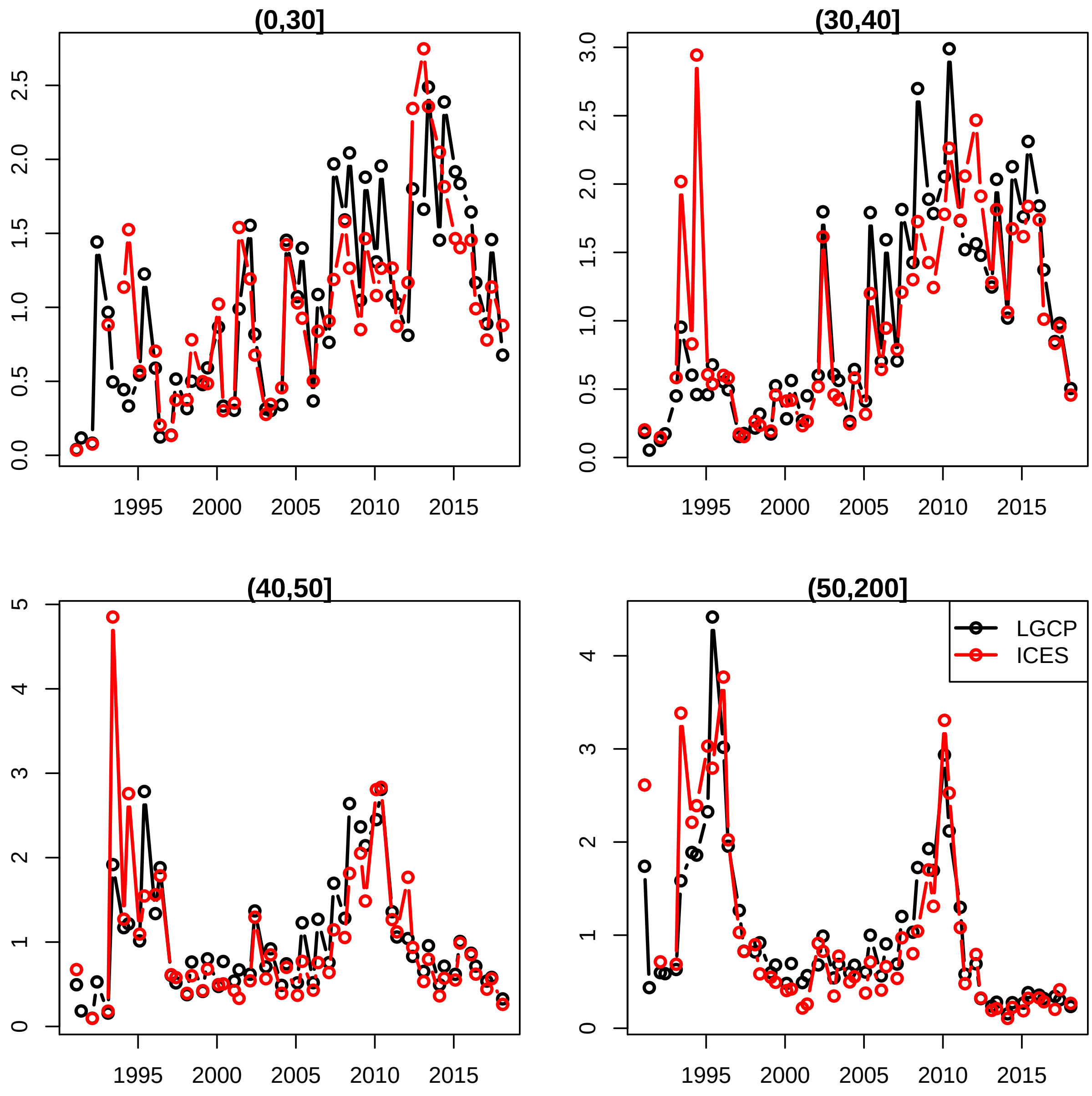


Figure 5: As figure 4 but for selected subsets of size groups.

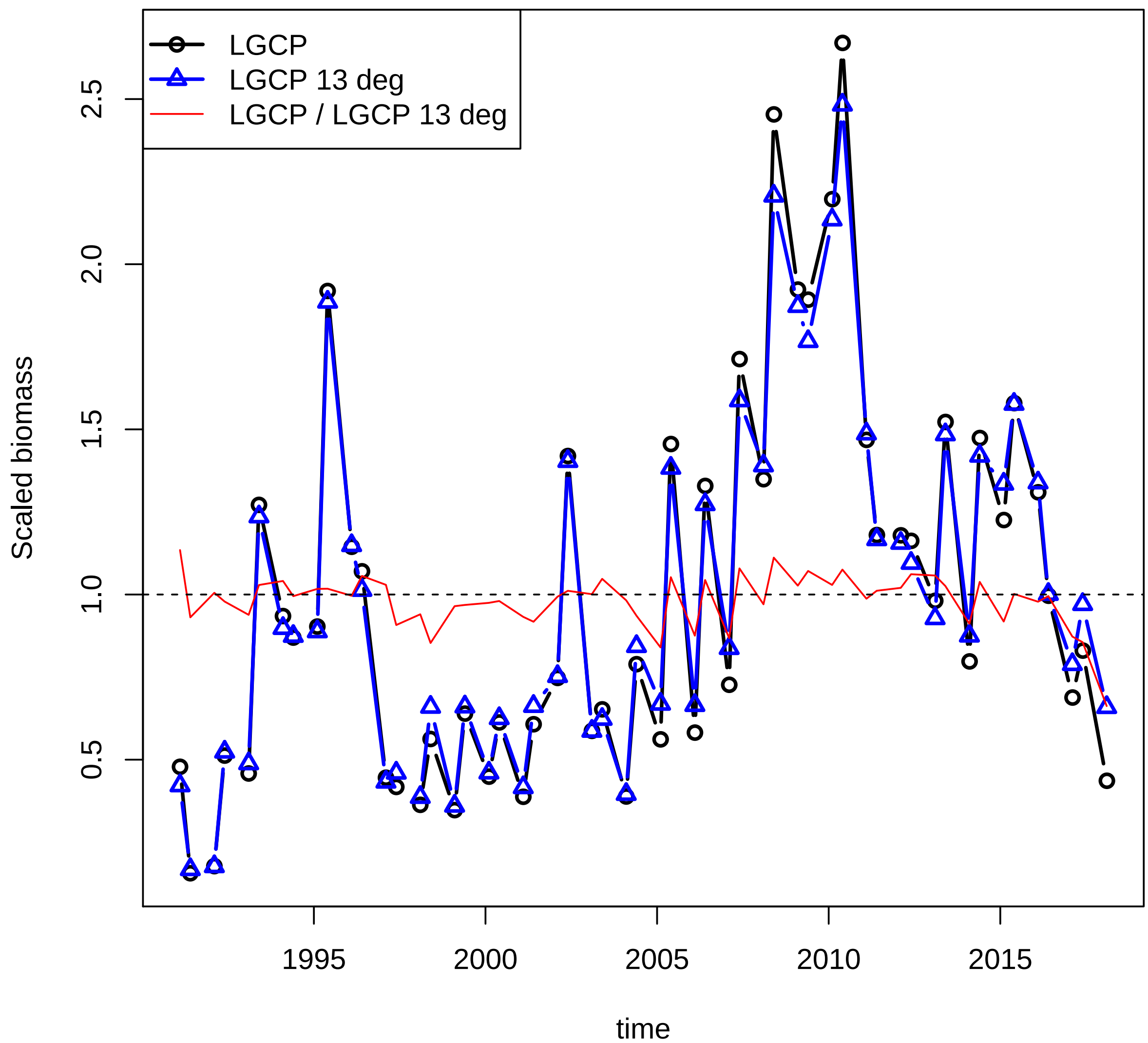


Figure 6: As figure 4, but comparing LGCP indices using the standard EBcod area (SD25+) with an alternative assessment area (east of 13 degrees longitude).

References

[1] Baltic Sea Hydrographic Commission. Baltic Sea Bathymetry Database version 0.9.3, downloaded from http://data.bshc.pro/ on 02-01-2018, 2013.

[2] Karin Hüssy, Krzysztof Radtke, Maris Plikshs, Rainer Oeberst, Tatjana Baranova, Uwe Krumme, Rajlie Sjöberg, Yvonne Walther, and Henrik Mosegaard. Challenging ICES age estimation protocols: lessons learned from the eastern Baltic cod stock. *ICES Journal of Marine Science*, 73(9):2138–2149, 2016.

[3] K Kristensen, A Nielsen, CW Berg, H Skaug, and B Bell. TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software*, 70(5):1–21, 2016.

[4] Kasper Kristensen, Uffe Høgsbro Thygesen, Ken Haste Andersen, and Jan E Beyer. Estimating spatio-temporal dynamics of size-structured populations. *Canadian Journal of Fisheries and Aquatic Sciences*, 71(2):326–336, 2013.

5 Appendix

	1	4
1991	305	79
1992	203	50
1993	271	98
1994	277	91
1995	302	79
1996	349	92
1997	336	125
1998	370	99
1999	378	216
2000	369	228
2001	345	186
2002	262	210
2003	297	212
2004	302	207
2005	317	262
2006	278	210
2007	275	229
2008	282	234
2009	298	227
2010	286	224
2011	294	241
2012	272	227
2013	304	206
2014	252	226
2015	275	214
2016	279	267
2017	306	276
2018	307	0

Table 1: Number of hauls by year and quarter

	21	22	23	24	25	26	27	28
1991	0	76	0	68	64	119	15	41
1992	1	66	2	72	41	47	9	14
1993	1	55	5	81	111	71	12	33
1994	0	57	5	80	105	84	14	23
1995	0	34	5	77	111	99	11	44
1996	6	46	5	81	107	146	12	38
1997	2	59	1	77	124	148	15	35
1998	1	63	1	77	149	143	10	25
1999	11	61	4	129	205	120	11	53
2000	10	56	4	123	178	121	14	91
2001	11	64	4	111	146	95	19	81
2002	13	48	7	101	130	94	19	60
2003	12	44	6	92	154	121	24	56
2004	11	41	7	94	162	131	14	49
2005	11	56	6	88	179	148	23	68
2006	12	48	6	85	156	113	18	50
2007	13	49	4	87	148	120	18	65
2008	13	68	5	81	143	116	22	68
2009	9	50	6	99	172	101	22	65
2010	12	58	6	80	149	124	20	61
2011	13	59	6	95	175	109	21	57
2012	13	48	6	100	173	86	20	53
2013	13	49	5	92	171	105	15	60
2014	11	53	6	99	167	70	15	57
2015	12	61	8	88	167	85	17	51
2016	13	72	10	105	168	114	18	46
2017	13	72	11	98	205	109	19	55
2018	6	38	5	52	109	65	5	27

Table 2: Number of hauls by year and ICES area

	FOT	GOV	GRT	H20	LBT	P20	SON	DT	TVS	HAK	TVL
1991	59	6	64	48	59	72	76	0	0	0	0
1992	10	29	57	78	0	33	46	0	0	0	0
1993	44	32	71	119	25	50	28	0	0	0	0
1994	62	15	69	122	0	72	28	0	0	0	0
1995	31	46	68	104	0	47	18	67	0	0	0
1996	77	6	55	111	0	70	11	85	26	0	0
1997	89	0	49	79	20	141	20	0	22	41	0
1998	81	0	65	114	0	121	22	0	23	43	0
1999	52	39	59	114	39	61	10	0	72	40	108
2000	16	70	82	98	59	52	0	0	101	37	82
2001	0	0	0	0	0	11	0	0	224	0	296
2002	0	0	0	0	0	0	0	0	211	0	261
2003	0	0	0	0	0	0	0	0	205	0	304
2004	12	0	0	0	0	0	0	0	211	0	286
2005	0	0	0	0	0	0	0	0	233	0	346
2006	0	0	0	0	0	0	0	0	163	0	325
2007	0	0	0	0	0	0	0	0	171	0	333
2008	0	0	0	0	0	0	0	0	186	0	330
2009	0	0	0	0	0	0	0	0	212	0	313
2010	0	0	0	0	0	0	0	0	173	0	337
2011	0	0	0	0	0	0	0	0	188	0	347
2012	0	0	0	0	0	0	0	0	182	0	317
2013	0	0	0	0	0	0	0	0	176	0	334
2014	0	0	0	0	0	0	0	0	224	0	254
2015	0	0	0	0	0	0	0	0	184	0	305
2016	0	0	0	0	0	0	0	0	204	0	342
2017	0	0	0	0	0	0	0	0	198	0	384
2018	0	0	0	0	0	0	0	0	98	0	209

Table 3: Number hauls by year and gear

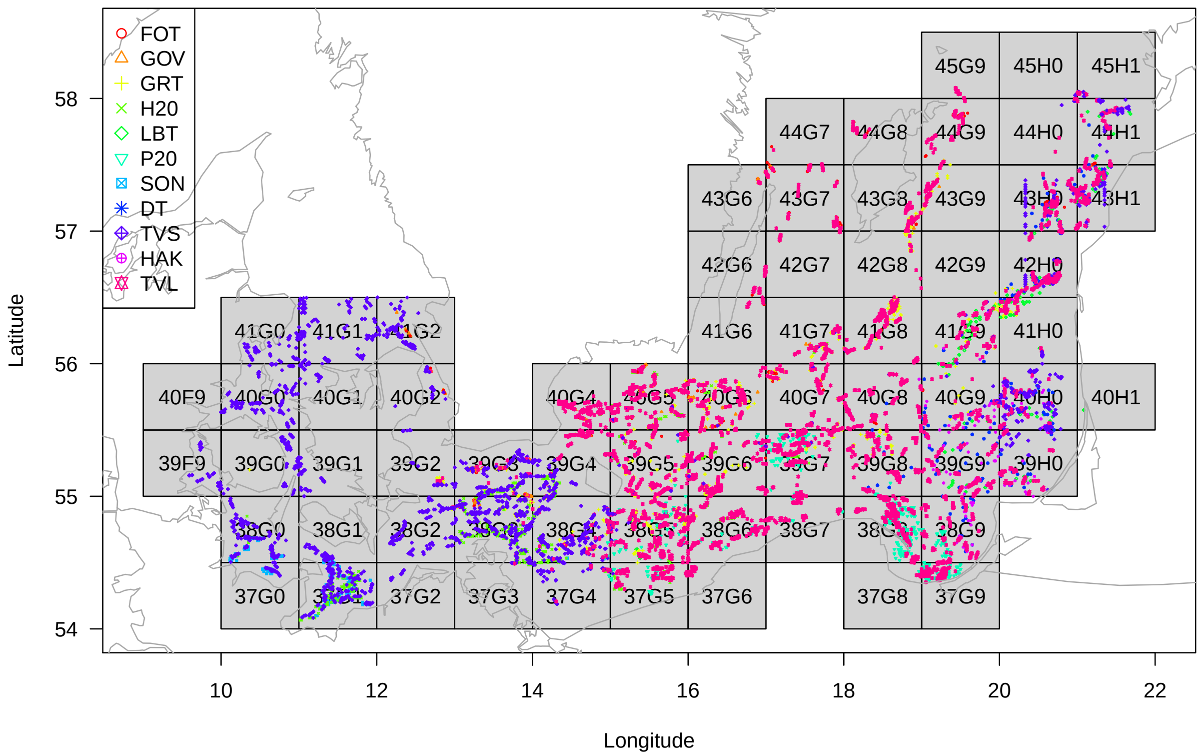


Figure 7: Map of all hauls (13402 in total) used by the model colored by gear type

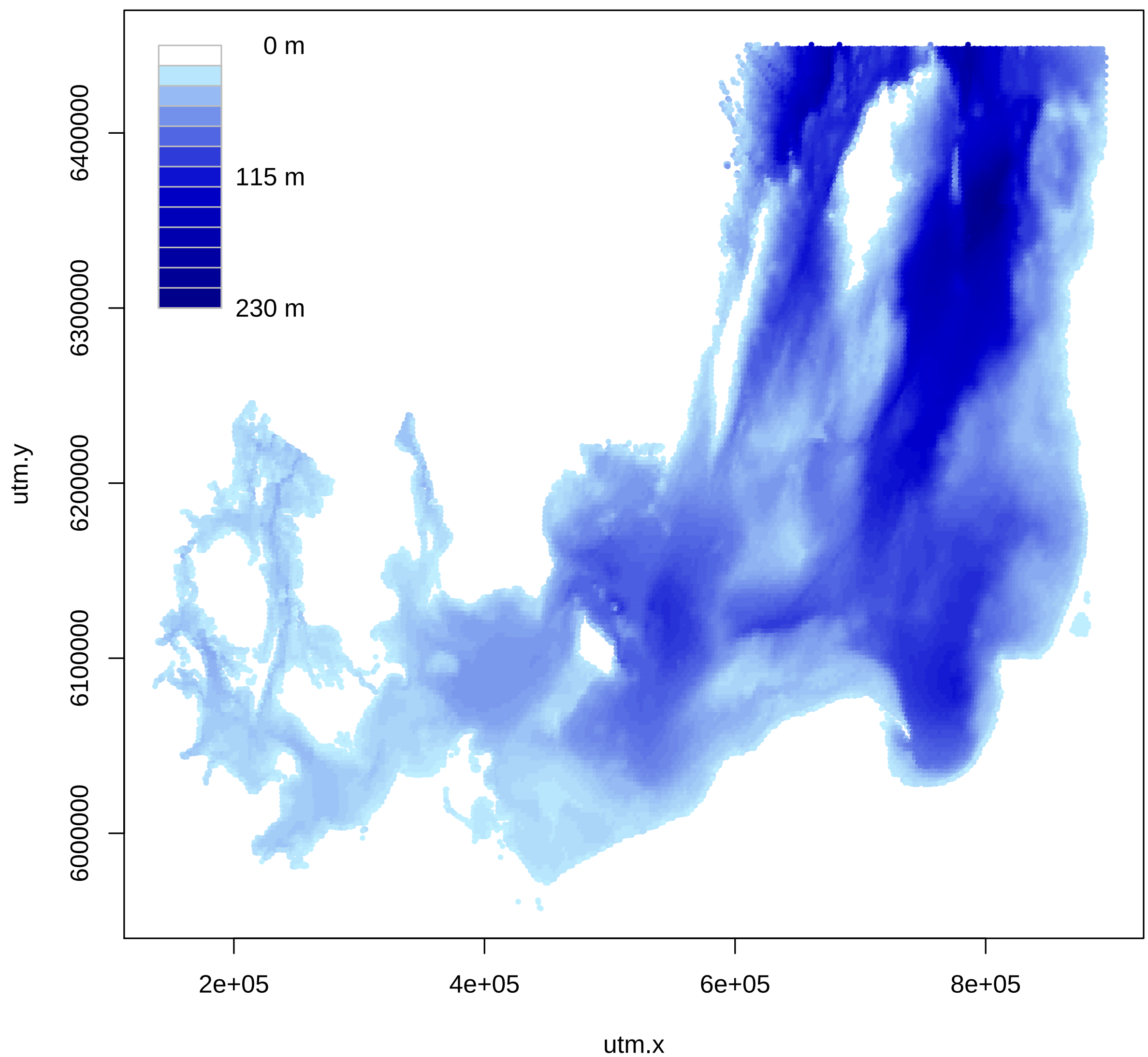


Figure 8: Bathymetric map (2500 m resolution) for the combined Eastern and Western Baltic cod areas.

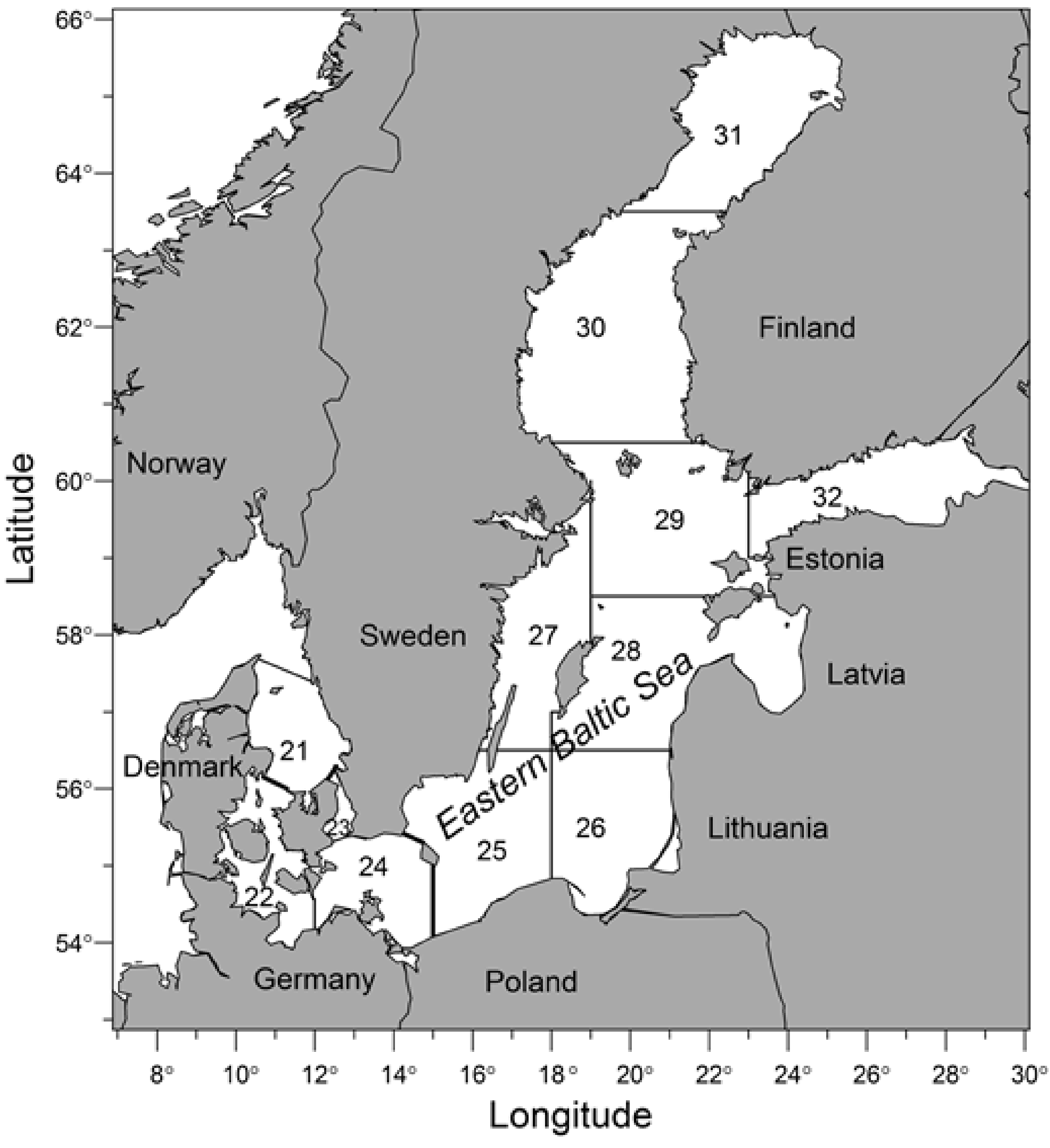


Figure 9: Map of ICES areas (source: [2])

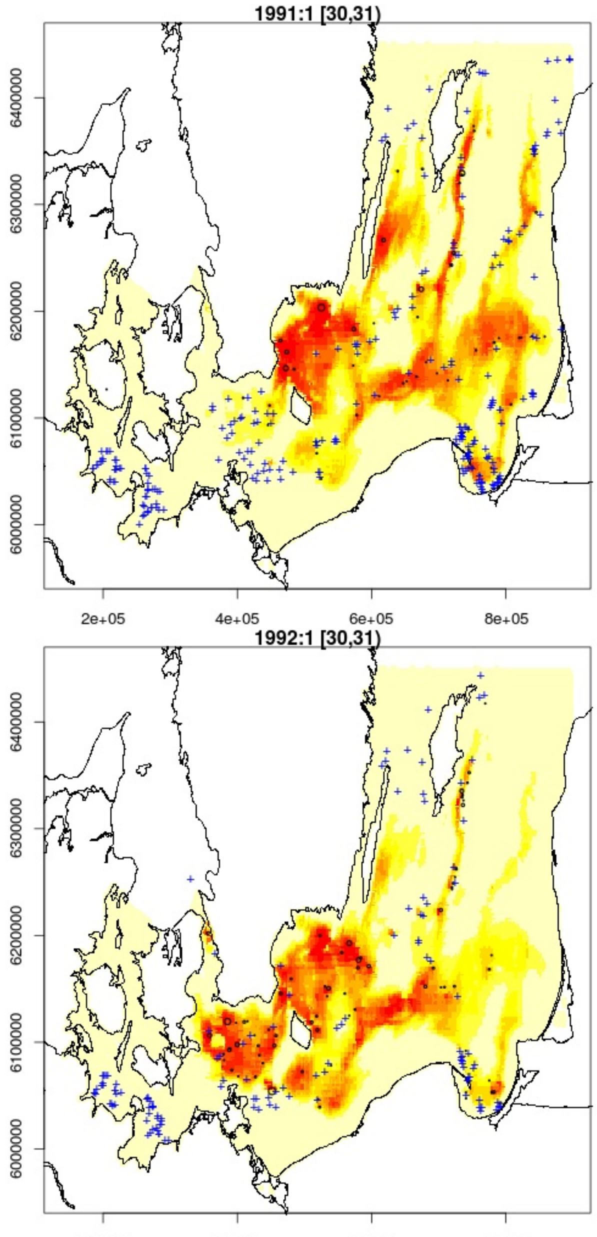


Figure 10: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

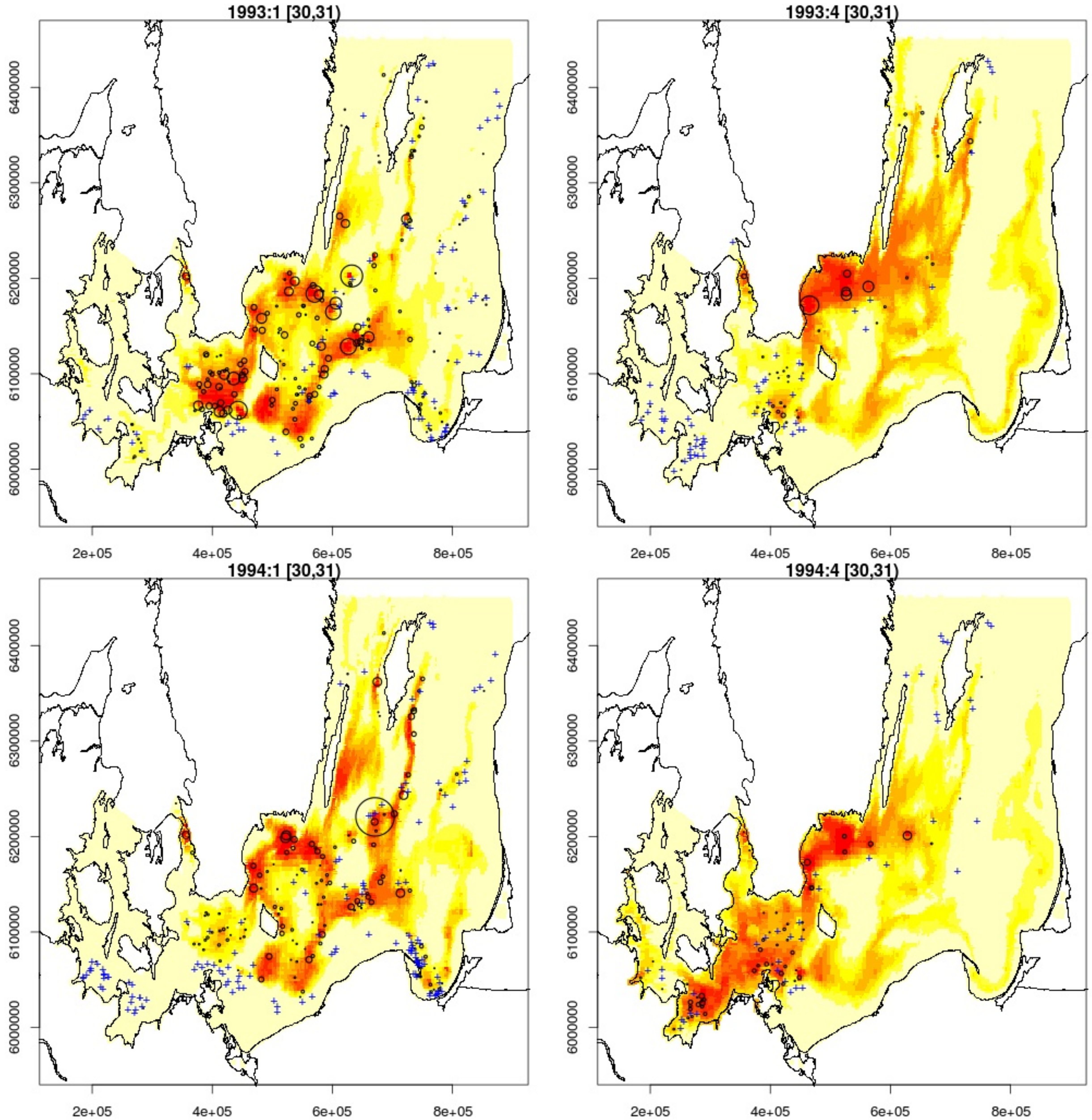


Figure 11: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

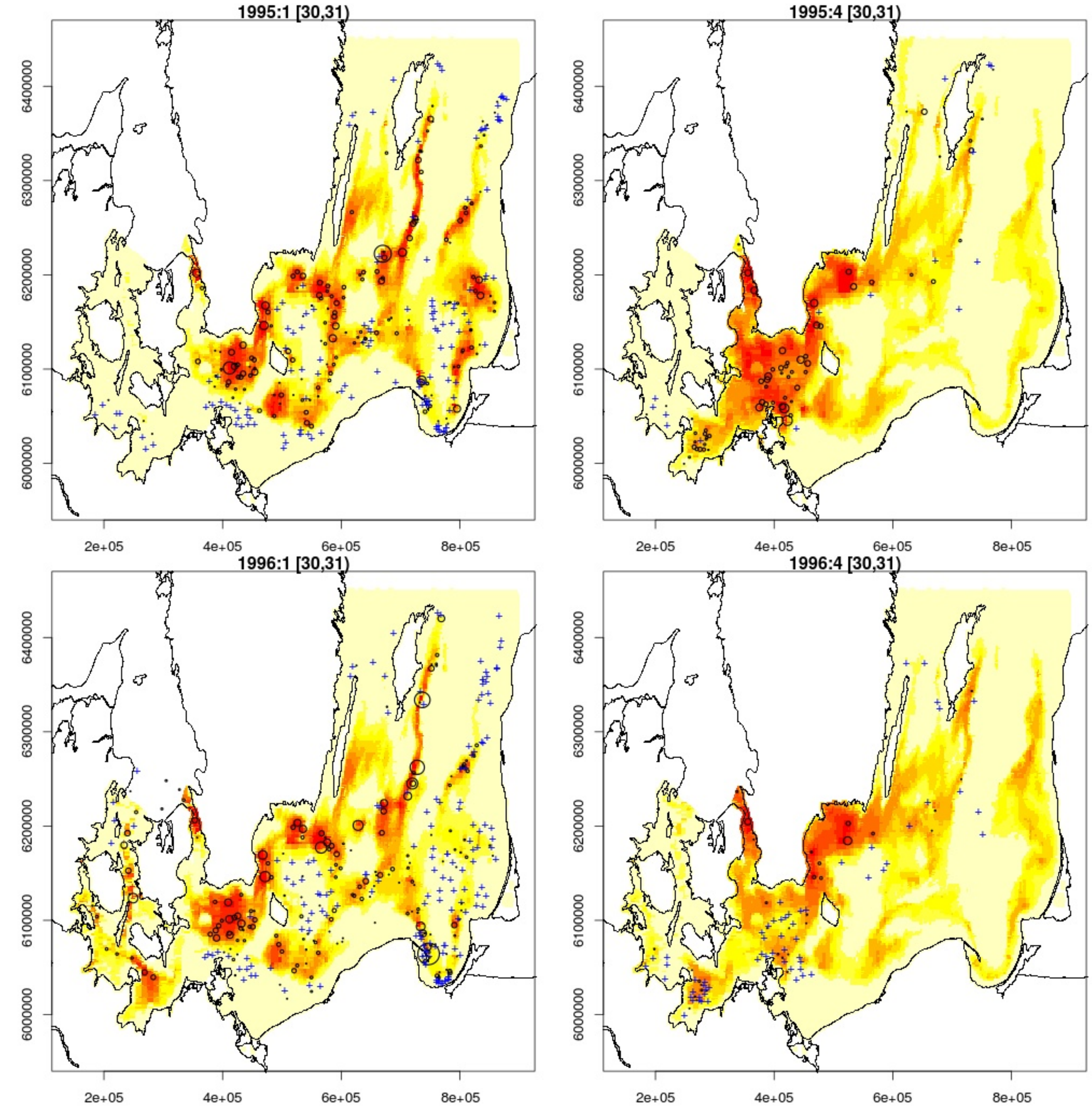


Figure 12: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

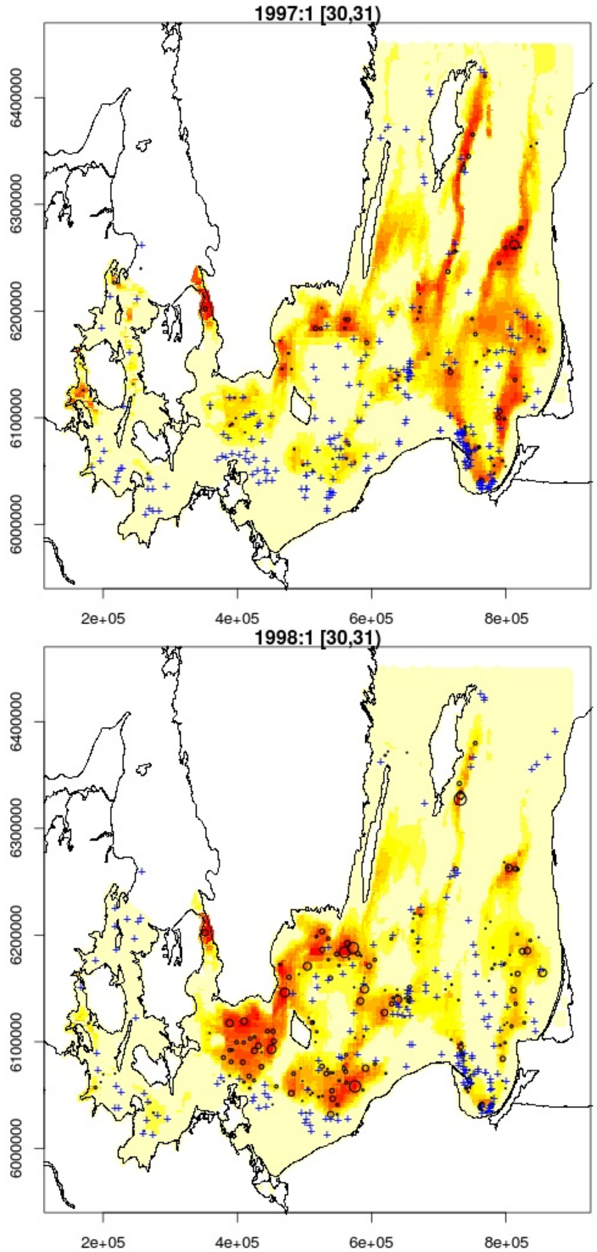


Figure 13: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

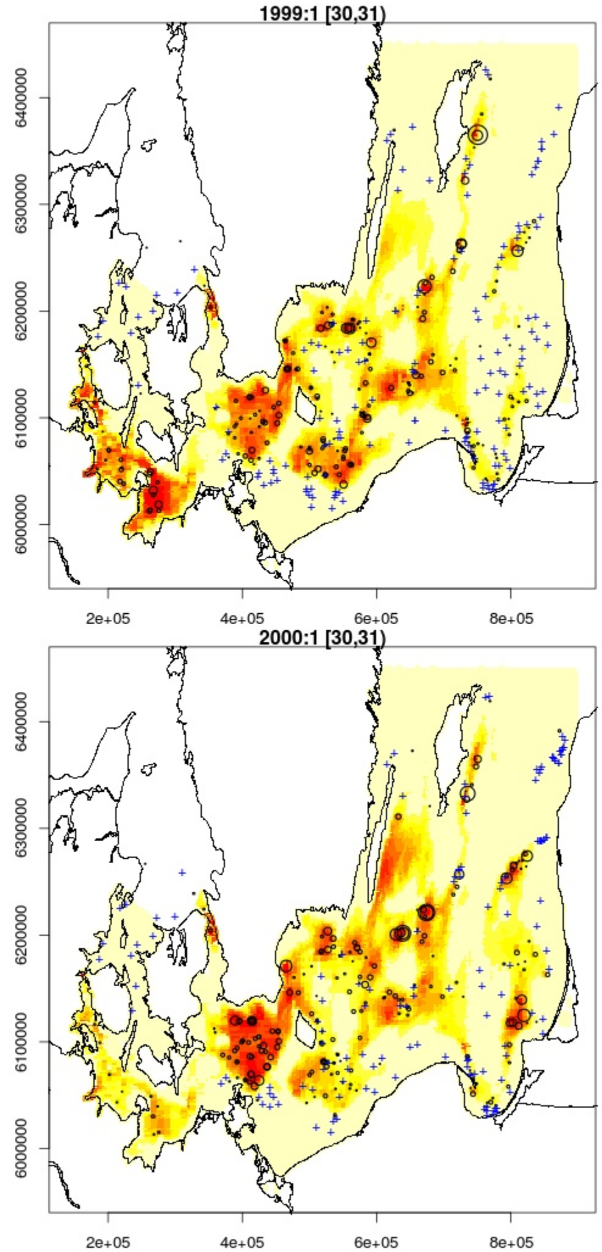


Figure 14: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

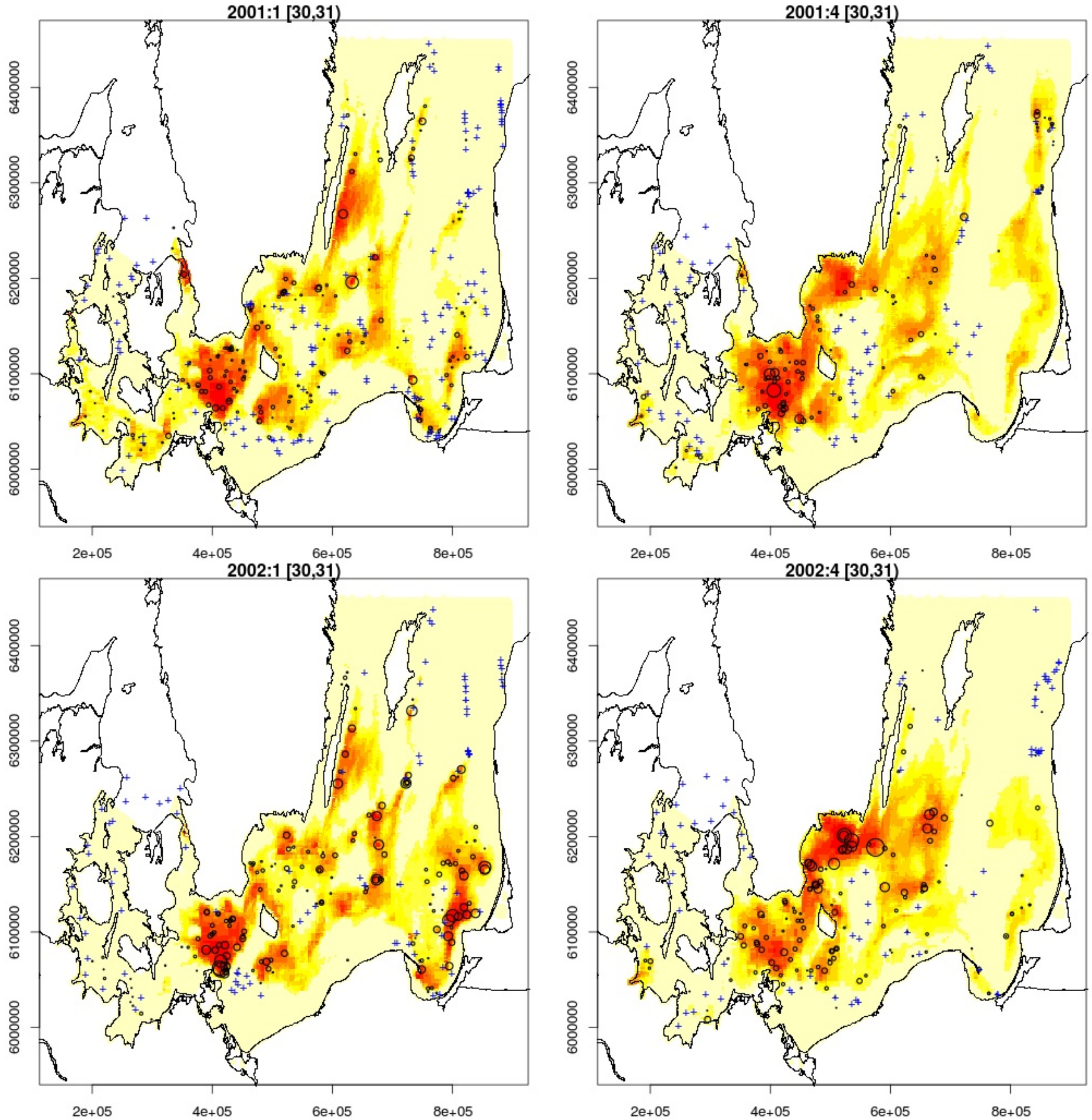


Figure 15: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

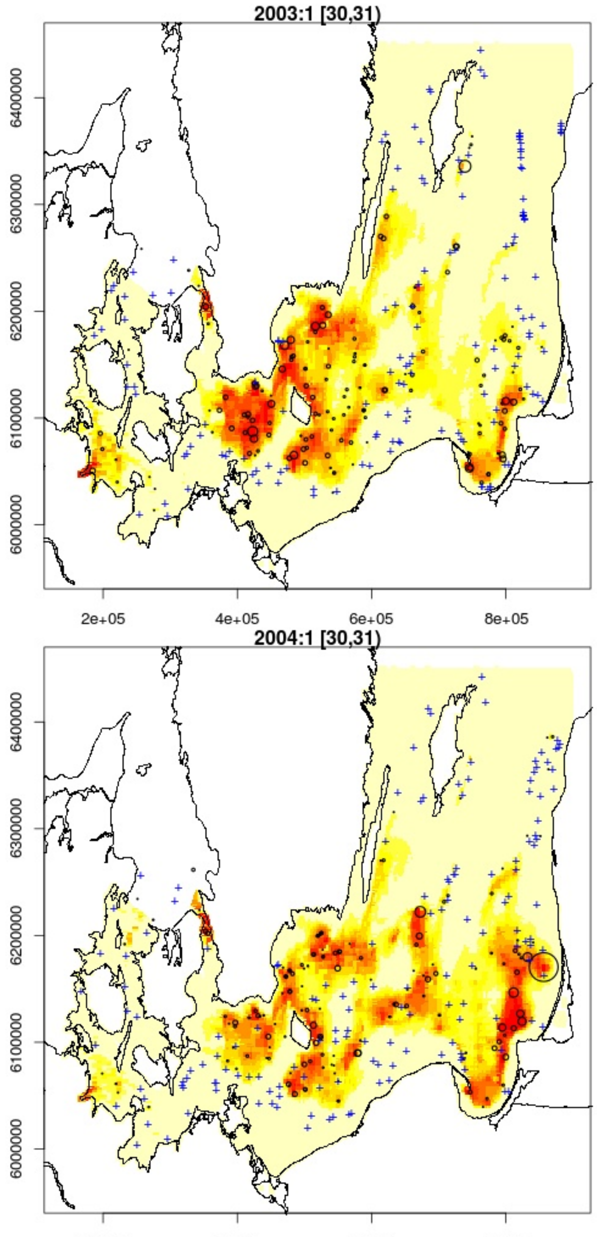


Figure 16: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

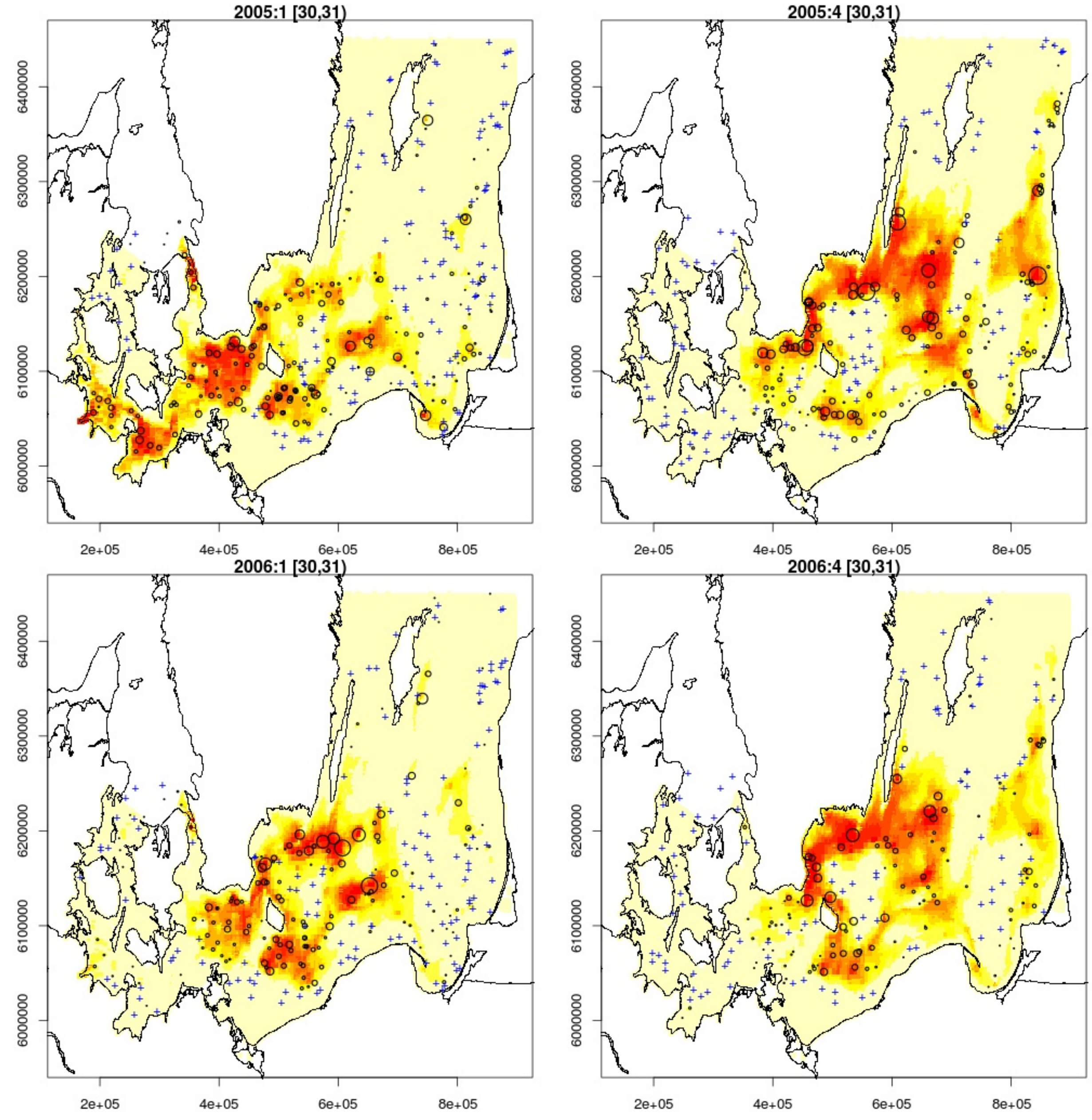


Figure 17: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

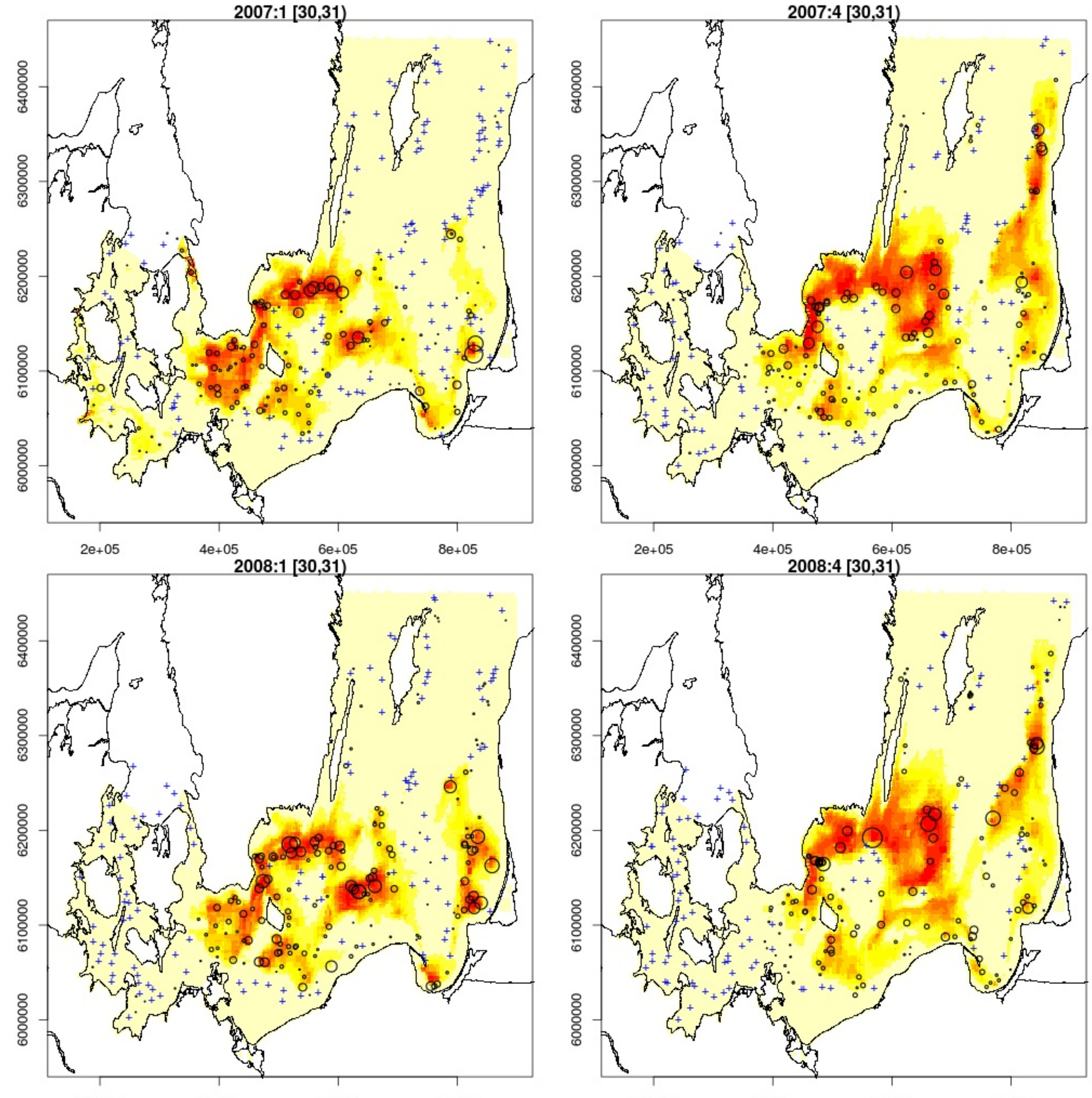


Figure 18: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

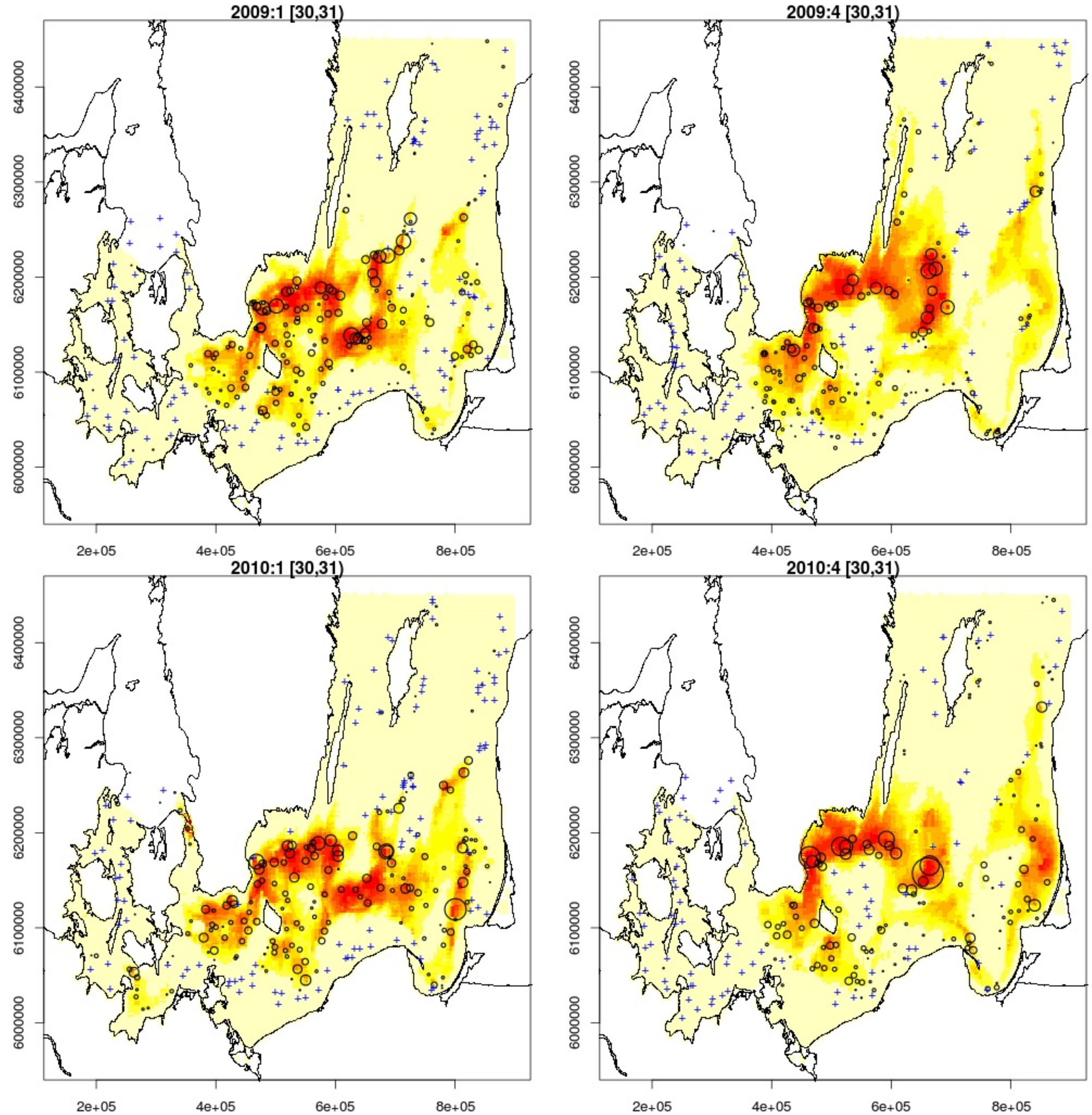


Figure 19: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

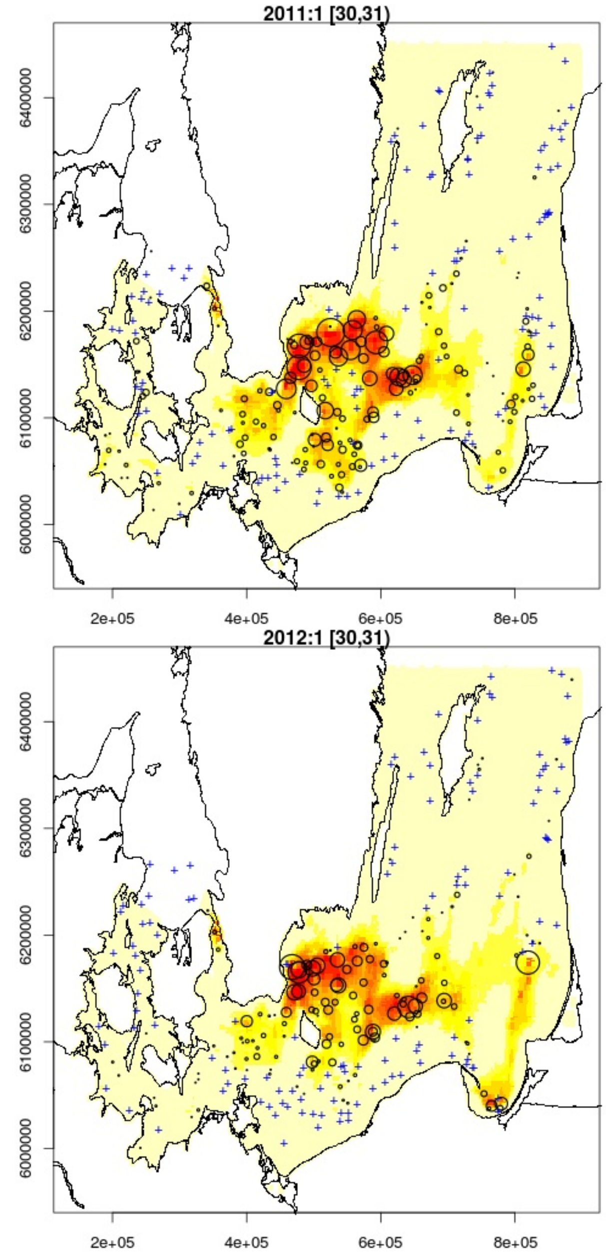


Figure 20: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

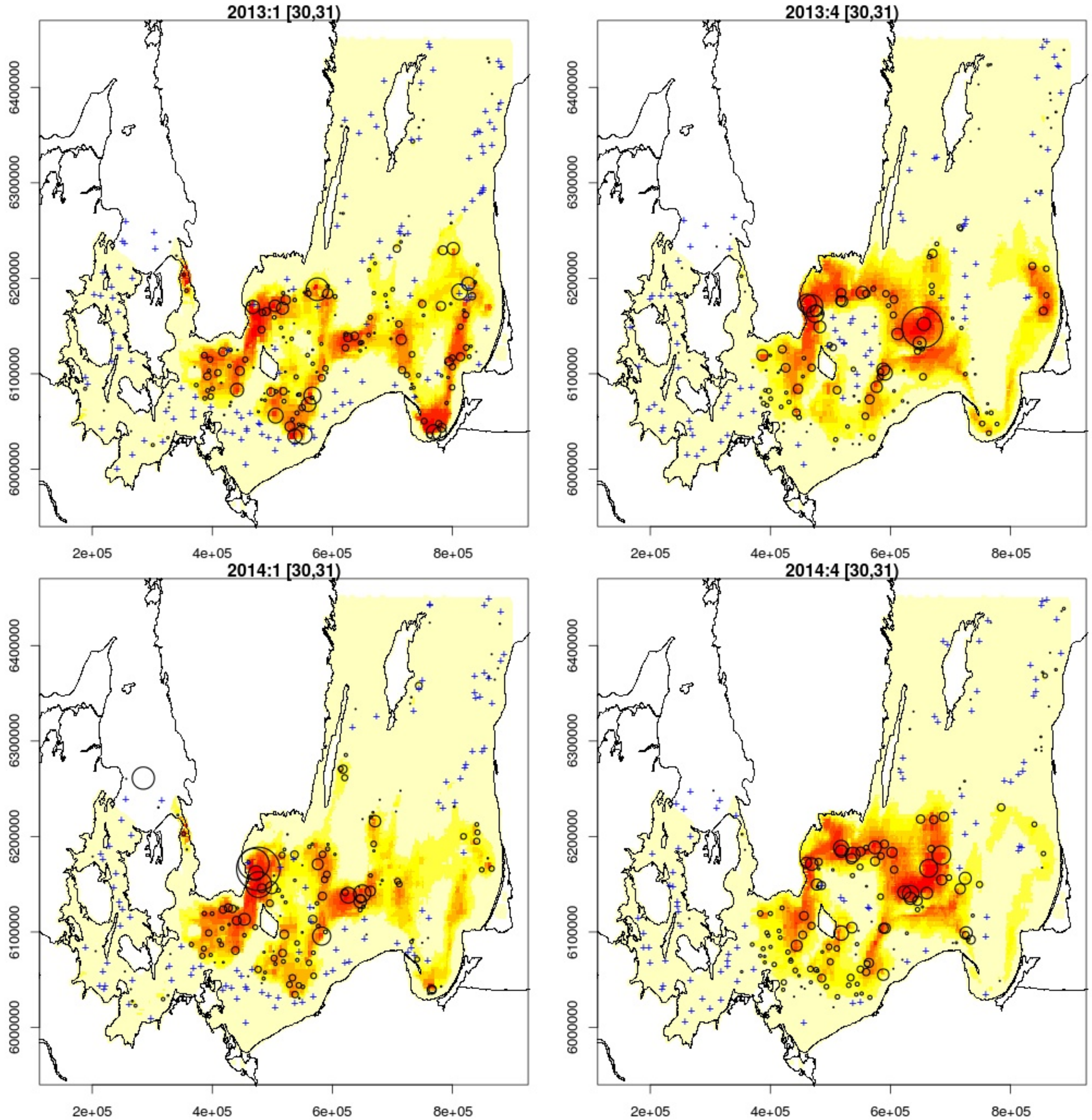


Figure 21: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

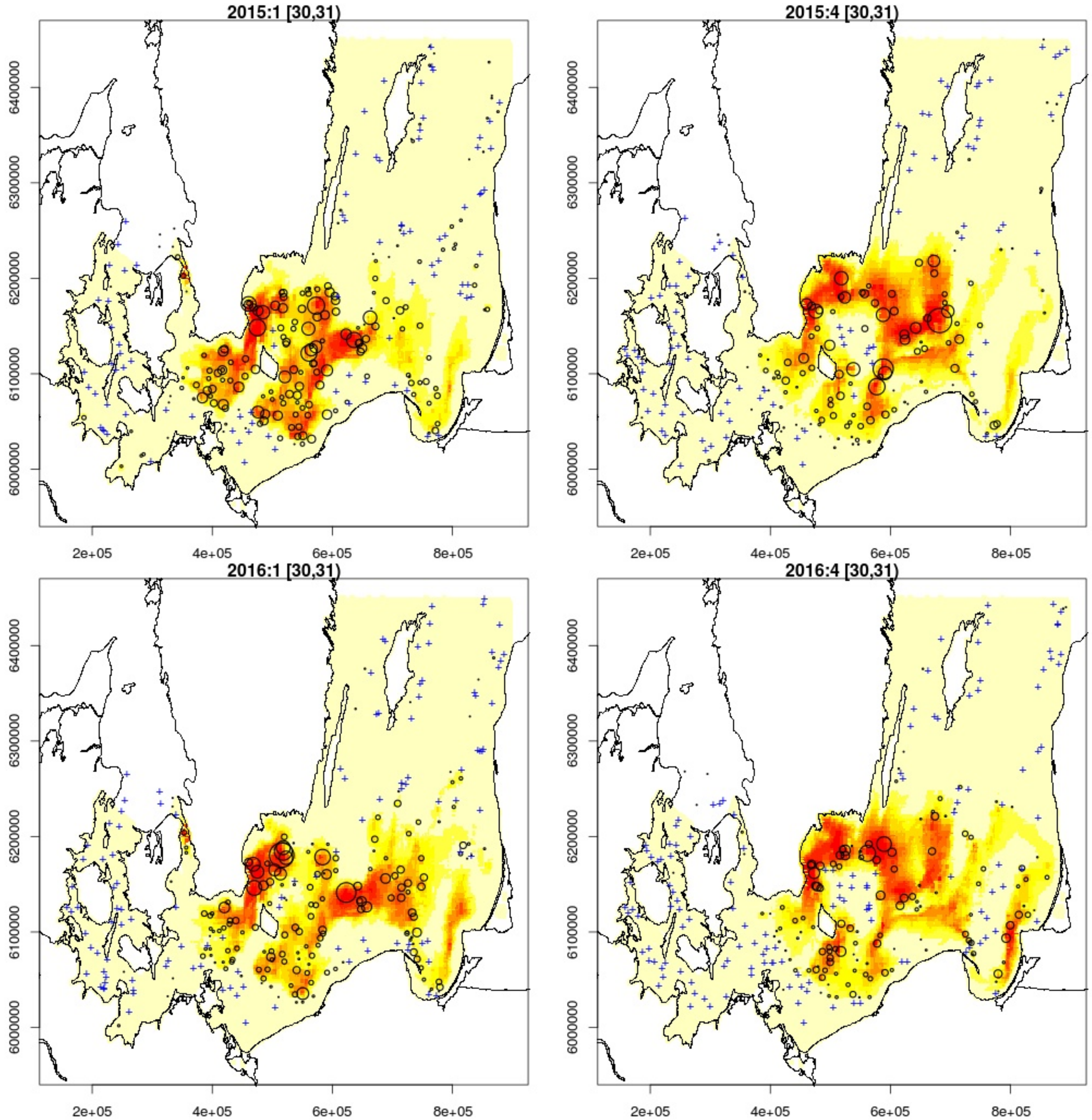


Figure 22: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

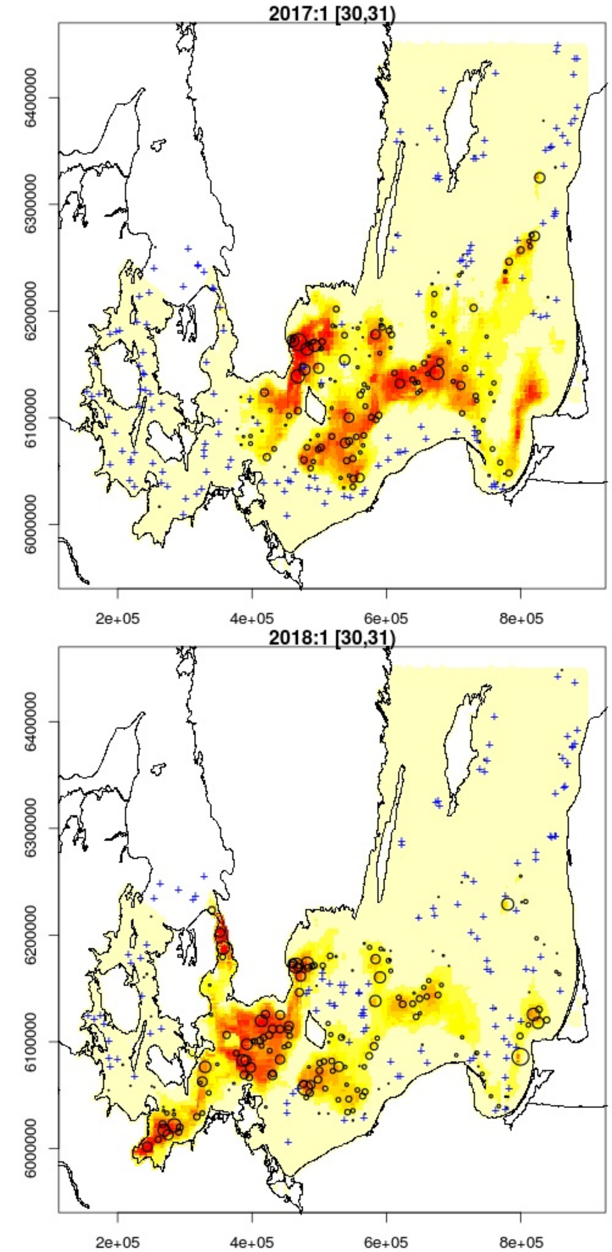


Figure 23: Concentration map (relative spatial abundance within a year) for the most frequently observed length group (30 cm). Hauls with positive catch rates are shown as circles with an area proportional to the numbers caught. Hauls with no catches are shown as blue plus signs.

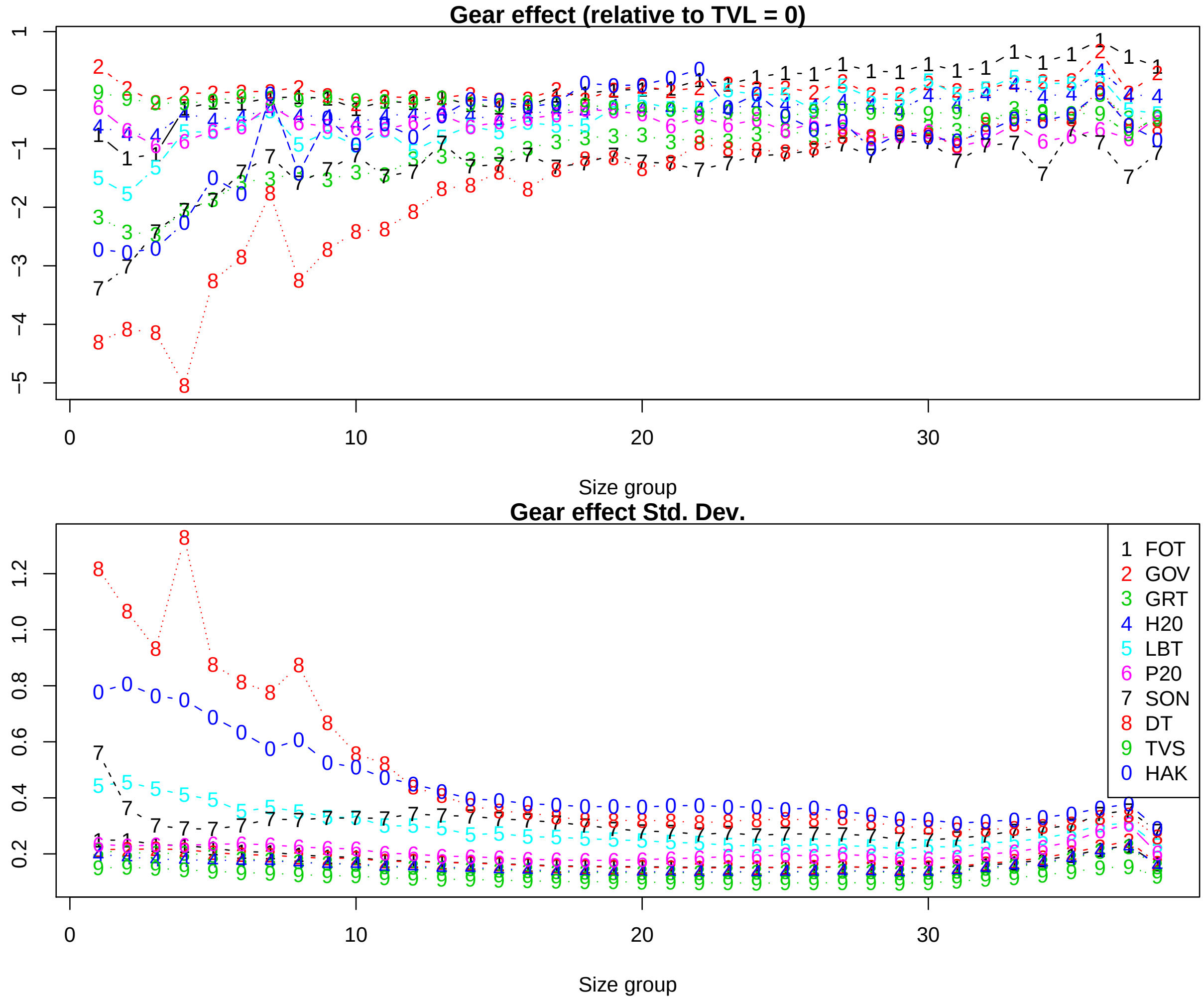


Figure 24: Estimated gear effect (log-scale) by length group.

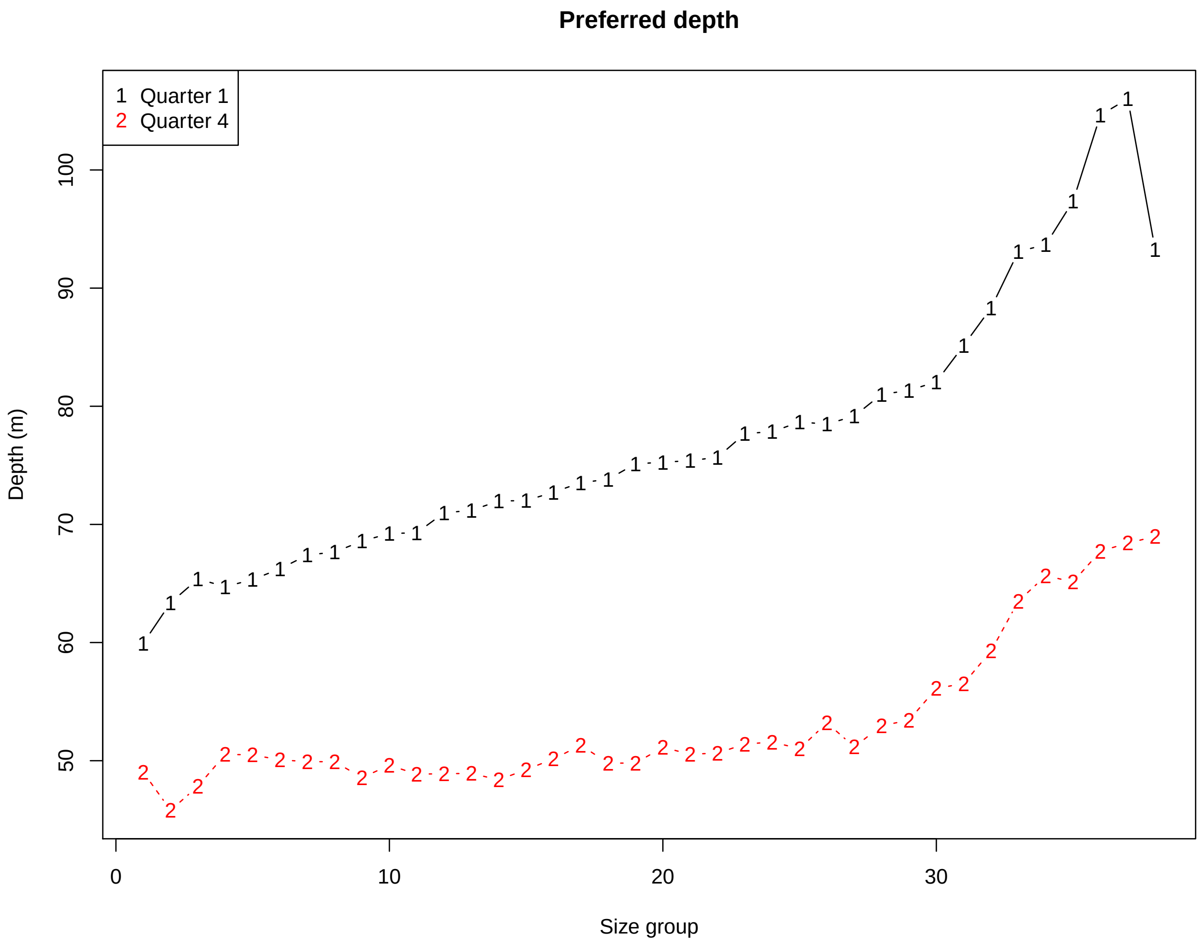


Figure 25: Estimated preferred depth by length group.

WD 3 - Application of the egg production method to estimate stock trends and spawning stock biomass of Eastern Baltic cod

Fritz Köster, Bastian Huwer, Gerd Kraus, Rabea Diekmann, Margit Eero, Andrei Makarchouk, Serra Orey, Jan Dierking, Piotr Margonski, Jens Peter Herrmann, Jonna Tomkiewicz, Daniel Oesterwind, Paul Kotterba, Rüdiger Voss

Summary

Since 30 years, standardised ichthyoplankton surveys have been carried out by Denmark, Germany, Latvia and Poland. Here, using this long-term data series, we deployed the annual egg production (AEPM) and the daily egg production method (DEPM) for the Bornholm Basin, the main spawning area of EB cod. The AEPM requires full egg survey coverage of the spawning season to estimate the annual egg production, and in addition relative fecundity and sex ratios to derive estimates of spawning stock biomass (SSB). The DEPM requires estimates of daily egg production at peak spawning time, individual spawning frequency, i.e. how many females participate in spawning at given dates as well as relative fecundity and sex ratios. Both methods yielded comparable results, and confirmed stock trends from Baltic International Bottom Trawl Surveys. For the years when reliable stock assessment is available, EPM yielded a similar trend in SSBs, being however consistently lower in absolute terms, especially at low stock sizes. Lower SSB than estimated in stock assessments can be explained by different processes, which can at least partly be taken into account when applying EPM. Egg production method (EPM) based estimates of stock trends and size can supplement the information from standard trawl surveys.

1 Data and methods

1.1 Egg abundance

Ichthyoplankton surveys have been carried out in the central and eastern Baltic regularly since the 1950's. A detailed description of surveys is given in ICES (2018 WGALES). In the present analysis, the abundance of cod eggs was determined for spring/early summer (May/June) and summer spawning (July/August) in different central and eastern Baltic basins in order to ascertain the importance of the Bornholm Basin as the only prominent spawning area (Köster et al., 2017). The ichthyoplankton gear in use was either a Bongo-net (60 cm diameter) equipped with 335 and 500µm mesh sizes, towed in a double oblique haul integrating the entire water column down to 2 m above the bottom or an IKS-80 net (80 cm diameter with 500 µm mesh size), vertically deployed through the water column or from 150m water depths upwards in deep water areas of the Gotland Basin (Fig. 1). A Bongo was consistently used in the Bornholm Basin and an IKS-80 in the Gdansk Deep and Gotland Basin until mid 1990's, afterwards a Bongo was deployed as well on certain cruises.

For the application of the Egg Production Method (EPM) in the Bornholm Basin, the youngest cod egg stage IA (Thompson and Riley, 1981) was determined on 3-14 ichthyoplankton cruises per year from 1991 to 2017. Sampling covered the spawning area of approximately 10800 km^2 (area enclosed by the 60 m isobath) on a regularly spaced station grid with approximately 10 nm miles grid-point distance. Until 1994 a station grid of 36 stations was covered, providing a slightly weaker coverage towards the boundaries of the spawning ground compared to later years where 45 stations were sampled.

All egg counts were standardised to $1m^2$ surface area by the volume of water filtered and the maximum depth of the tow and raised to the entire Bornholm Basin (area with water depths >60m), Gdansk Deep (water depths >70m) and Gotland Basin (water depths >80m).

1.2 Egg production

For each survey a single estimate of the Bornholm Basin wide daily egg production of the youngest egg stage IA was calculated. The duration of a survey was on average three days and the median date of each survey was assumed to be representative for each production estimate. Daily egg production was calculated as total abundance of the youngest egg stage IA divided by the stage duration estimated from stage-specific egg development–temperature relationships (Wieland, 1995). The ambient temperatures used in the relationship was derived from the ICES hydrographic database for Sub-division 25 in 2nd and 3rd quarter, respectively, applying a model to predict the relative vertical distribution of cod eggs until 1999. In this model, empirical vertical egg distributions are related to water density profiles by fitting a parabolic function to the log relative distribution (Köster et al., 2005). For the period since 2000, the average temperature in 60-80 m depths, i.e. the water layer comprising highest egg abundances (Wieland et al., 1997), were extracted from the ICES hydrographic database for Sub-division 25 in 2nd and 3rd quarter, respectively.

Dead eggs that could not be assigned to a specific stage were distributed proportionally to the relative stage-specific distribution of live eggs in the samples to estimate the total realised egg production (total amount of spawned eggs) rather than the viable production. None fertilised eggs were considered to be stage IA unless the decay process did not allow the identification of no development beyond the one cell stag. In that case they were treated as dead eggs. A correction for mortality within the egg stage IA was applied, with the mortality estimated from egg stage IA and IB daily egg production estimates for each survey, averaged over surveys, however considering seasonal differences in egg mortality by calculating average mortality rates for the 1st and 2nd part of the spawning season.

The total seasonal egg production of cod in the Bornholm Basin was determined from an annual production curve of egg stage IA derived by fitting GLM's to survey specific daily production values assuming a negative binomial distribution in spawning activity. The first survey was usually performed in March, the last survey in late summer or in November (since 2008), due to the protracted spawning season of EB cod (Wieland et al., 2000). In order to model daily egg production during the course of the year, it was necessary to fill survey gaps at the beginning or the end of the spawning season. Both times are generally characterised by zero to very low egg numbers. Therefore, a zero egg production until mid January was assumed, and if the end point of the spawning season could not be adequately covered by a survey, linear interpolation between November surveys in preceding and subsequent years were applied.

According to Daan (1981), daily cod egg production of the youngest egg stage in a given year is normally distributed over time. Following this rational we modelled the total annual egg production from survey specific daily egg production with a generalised linear model (GLM) using a log link and a linear predictor of the form:

$g(x) = a + b*Day + c*Day^2$

Errors were assumed to be Poisson distributed, but due to high overdispersion the negative binomial distribution was finally applied. The resulting curves follow a Gaussian shape and are symmetric around the annual specific date of peak spawning. The model was validated graphically according to the model fit (predicted values and deviance residuals vs. fitted values) and the existence of influential observations. Annual egg production was then calculated by integrating the area below the curve for each single year. All calculations were performed using R (version 3.5.1) as language for statistical computing (R Core Team, 2012).

To validate the robustness of the seasonal egg production against assumptions on start and endpoint of the spawning season, the daily egg production at peak spawning time was determined and used in an alternative approach to estimate the stock biomass not only by the Annual egg production method (AEPM) but also by the Daily egg production method (DEPM).

1.3 Individual fecundity and stock structure

Annual mean values of relative potential fecundity (numbers of developing vitellogenic follicles per gram female body weight during final maturation) in ICES Sub-division25 were obtained from Kraus et al. (2002) for 1991, 1992, 1996 and 1998-2000, STORE (2003) for 1995 and Örey (2018) for 2005, 2007, 2009, 2011, 2013 and 2016. In case of lacking annual mean values, an average between preceding and subsequent years or a linear interpolation over time was applied. As a the relative fecundity in most recent years showed a length dependence, year specific average relative fecundity values were derived from weighting length specific values with length stock structure data from the 1st quarter International Bottom Trawl Survey (BITS). No correction for atresia was applied following the methodology and reasoning of Kraus et al. (2012). The female ratio in the stock was assumed to be 50%.

1.4 Estimation of spawning stock biomass

In the application of the AEPM, the spawning stock biomass (SSB) of cod in the Bornholm Basin was estimated by:

$$SSB = AEP \;/\; (\; s \;*\; Frel \;)$$

AEP = Total production of egg stage IA from seasonal egg production curves

s = Sex ratio, i.e. proportion of females

Frel = Relative fecundity, i.e. average numbers of developing vitellogenic follicles per gram female body weight

As an alternative to AEPM the spawning stock biomass of cod in the Bornholm Basin at peak spawning time was estimated applying an adaptation of the DEPM described by Lasker (1985). Average values for spawning frequency and individual spawning time were applied as data to estimate yearly specific values were not present for all years considered in the analyses.

$$SSB_t = DEP_t \;/\; (s_t \;*\; Frel) \;*\; D$$

DEP_t = Daily production of egg stage IA during peak spawning time

s_t = Sex ratio, i.e. proportion of females during peak spawning time

Frel = Relative fecundity

D = Individual spawning season in days was assumed to be 75 days, as determined by Tomkiewicz and Köster, 1999).

Time trends in SBB estimates are compared with catch rates of adult EB cod from 1st and 4th quarter Baltic International Trawl Surveys (ICES 2018) as well as results from the last accepted analytical assessment (ICES 2013). Adult cod in the survey were defined as fish ≥30 cm until 2010 and ≥25 cm from 2011 onwards to account for the reduction in average size at reaching sexual maturity (e.g. Eero et al., 2015).

2 Results

2.1 Egg abundance

Time series of egg abundances during May/June and July/August from different eastern Baltic basins, confirm the Bornholm Basin as the most important spawning area of the EB cod stock since 1985 (Fig. 2), which is in contrast to earlier decades (Karasiova et al., 2009; Köster et al., 2009). Egg abundances were increasing in most recent years early in the spawning season (May/June) in both the Gdansk Deep and the Gotland Basin, while a similar increase is not evident for summer month (Fig. 2). However, the survey coverage of the Gotland Basin was limited in most recent years. A relatively high egg abundance was encountered in the Gdansk Deep in summer 2003, being one of the years with a major Baltic inflow (e.g. Morholz et al., 2015). Above average egg abundances were, however, not detected in the Gotland Basin and despite still improved hydrographic conditions in 2004 in neither the Gdansk Deep and the Gotland Basin.

Considering years in which all basins are sufficiently covered, revealed that on average 14,8% of the total egg abundance occurred in the Gdansk Deep and Gotland Basin. Thus, estimating the spawning stock biomass (SSB) based on ichtyoplankton survey conducted in the Bornholm Basin only, will result in an underestimation of the SSB in the same order of magnitude.

Egg abundance in the Bornholm Basin showed substantial interannual variability with high abundances in mid 1990'ties, followed by historic low values comparable to the early 1990's (Fig. 2). until increasing again after 2000 (May/June) and 2004 (July/August).

2.2 Egg production

The annual egg production (AEP) and the daily egg production (DEP) at peak spawning follow the same trend, being high from 1994-1997, low in 1998-2004 and subsequently increasing until 2010 (Fig. 3). From 2011-2013, egg production is reduced, to a level which is continued according to the AEP, while DEP suggests an increase in most recent years again. This relatively high deviation between both methods in most recent years is caused by a reduced spawning season with high egg production at peak spawning time and relatively low production early and late in the spawning season. In contrast, in 2006-2009 the egg production is relatively high over an extended spawning period, resulting in lower DEP compared to AEP, especially in 2008.

2.3 Individual fecundity

The applied fecundity time series indicates an upward trend with nearly doubling of the average annual relative fecundity values from 1991-2016 (Fig. 4). The most recent analyses of EB cod fecundity (Örey 2018) revealed for the period 2005-2011 average values of 769-802 eggs per g body weight, which corresponds well to values determined by Kraus et al. (2002) for 1996-2000, ranging from 780-892 eggs per g body weight, being higher than values determined by Kraus et al. (2002) for the period 1991-1992.

Data from Örey (2018) indicate a higher relative fecundity in 2013 and 2016 amounting to 988 and 918 eggs per g body weight, respectively (Fig. 4).

2.4 Spawning stock biomass

Application of AEPM derived two pronounced peaks in spawning stock biomass (SSB) of 80-90,000 t in 1994 and 2009, respectively (Fig. 5). SSB values below 20,000 t were determined for 1991/1992 and 1998-2003 followed by a steady increase until 2009 and a subsequent decline to 23,000 t in 2013, partly recovering afterwards to 31-42,000 t. The DEPM derived SSB estimates are quite similar, but somewhat higher in 1994, reaching 112,000 t and somewhat lower in the 2nd half of the 2000's reaching 64,000 t in 2010 (Fig. 5). SSB values in most recent years were slightly higher than determined by AEPM, ranging from 45-60,000 t.

Results followed the large scale stock trends of the BITS surveys (Fig. 6). The 1. quarter survey showed a somewhat later spawning stock increase in the 2000's, while the 4. quarter survey confirmed the EPM trend. The 1. quarter survey indicates as well as lower stock in mid 1990's than in the 2nd half of the 2000's.

Comparing the EMP derived absolute spawning stock estimates with results from the last accepted analytical assessment (ICES 2013) showed both similar trends and SSB's in the same order of magnitude, however the EMP derived estimates being consistently lower, apart from the period 2006-2009, when AEMP estimated values are similar to the assessment output (Fig. 7).

References

Daan, N., 1981. Comparison of estimates of egg production of the Southern Bight cod Stock from plankton surveys and market statistics. Rapp. P.-v. Réun. Cons. Int. Explor. Mer 178, 242-243.

ICES 2013. Report of the Baltic Fisheries Assessment Working Group (WGBFAS), 10-17 April 2013 ICES HQ, Copenhagen, Denmark, ICES CM 2013/ACOM:10.

ICES 2018. WGALES

ICES 2018. Benchmark data prep. report

Karasiova, E.M., Kraus, G., Voss, R., 2009. Long-term fluctuations of Baltic cod egg production in relation to major inflow events. ICES CM 2009/ E: 07.

Köster F.W., Möllmann C., Hinrichsen H.H., Wieland K., Tomkiewicz J., Kraus G., Voss R., Makarchouk A., MacKenzie B.R., St.John M.A., Schnack D., Rohlf N., Linkowski T., Beyer J.E. (2005) Baltic cod recruitment - the impact of climate variability on key processes. ICES Journal of Marine Science 62:1408-1425.

Köster, F.W., Vinther, M., MacKenzie, B.R., Eero, M., Plikshs, M. (2009) Environmental Effects on Recruitment and Implications for Biological Reference Points of Eastern Baltic Cod (Gadus morhua). Journal of Northwest Atlantic Fishery Science 41: 205–220.

Kraus, G., Tomkiewicz, J., Köster, F.W., 2002. Egg production of Baltic cod in relation to variable sex ratio, maturity and fecundity. Can. J. Fish. Aquat. Sci. 59, 1908-1920.

Kraus, G, Hinrichsen, H-H, Voss, R, Teschner, E, Tomkiewicz, J, Köster, F (2012). Robustness of egg production methods as a fishery independent alternative to assess the Eastern Baltic cod stock (*Gadus morhua callarias L.*). Fisheries Research 117: 75-85.

Mohrholz V., Naumann M., Nausch G., Krüger S., Gräwe U.(2015) Fresh oxygen for the Baltic Sea — An exceptional saline inflow after a decade of stagnation. Journal of Marine Systems 148: 152–166.

Örey S. (2018) Thinner females - fewer eggs? Temporal trends in Eastern Baltic cod fecundity (2005-2016). Master thesis GEOMAR, Helmholtz Centre for Ocean Research Kiel, 59 pp.

STORE, 2003. Environmental and fisheries influences on fish stock recruitment in the Baltic. Final Report to the EU Commission FAIR-98-3959. http://orbit.dtu.dk/en/projects/environmentaland-fisheries-influences-on-fish-stock-recruitment-in-the-balticsea-store (49b2c342-46ae-45fe-9a11-c0c43934d4bc).html.

Thompson, B.M., Riley, J.D., 1981. Egg and larval developmental studies in the North Sea cod (*Gadus morhua* L.). 1. Rapp. P.-v. Réun. Cons. Int. Explor. Mer 178, 553-559.

Tomkiewicz, J., Köster, F.W., 1999. Maturation processes and spawning time of cod in the Bornholm Basin of the Baltic Sea: preliminary results. ICES C.M. 1999/Y:25.

Wieland, K., 1995. Influence of hydrography on vertical distribution and egg mortality of the Baltic cod (*Gadus morhua callarias*) in the Bornholm Basin, southern central Baltic. Ph.D. thesis, University of Kiel, Ber. Inst. f. Meeresk. Kiel, 266, 114pp.

Wieland, K., and Jarre-Teichmann, A. 1997. Prediction of vertical distribution and ambient development temperature of Baltic cod (*Gadus morhua* L.) eggs. Fisheries Oceanography, 6: 172-183.

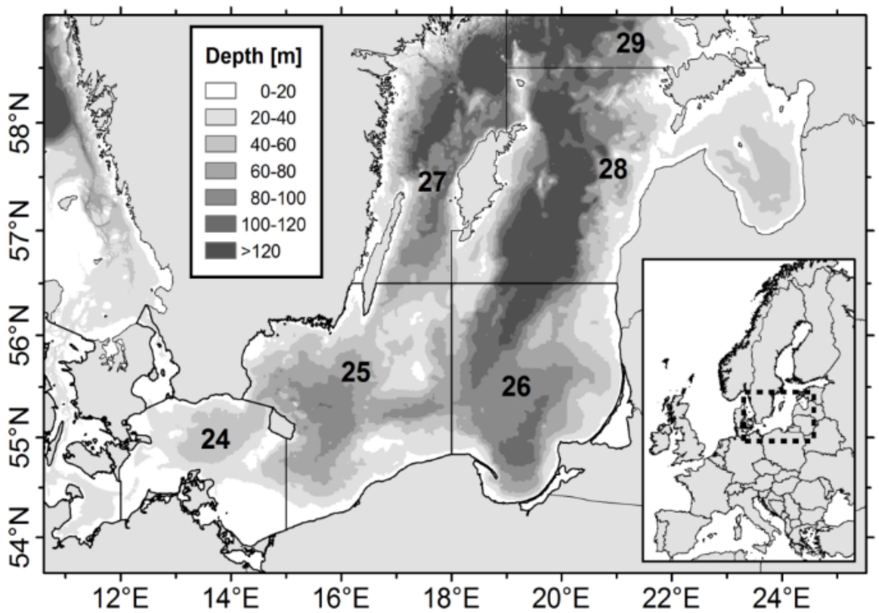
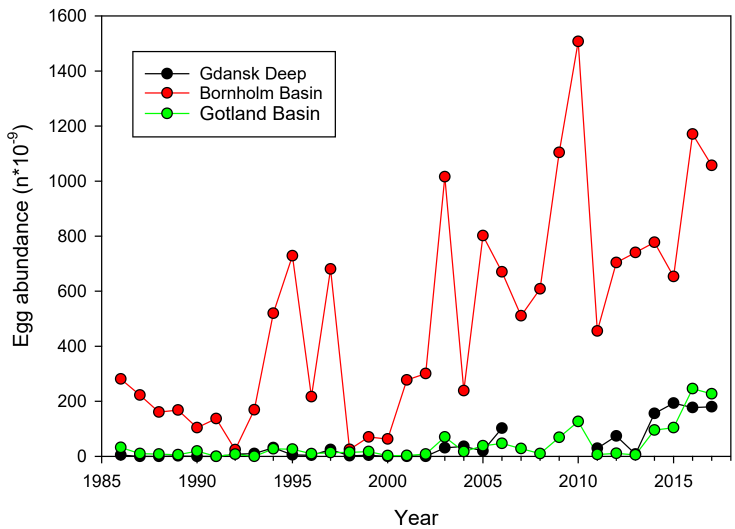


Figure 1. Study area in the central and eastern Baltic Sea including ICES Subdivisions (SD) as numbers and thin black lines; SD 24 corresponds to the wider area of the Arkona Basin, SD 25 to the Bornholm Basin, southern SD 26 to the Gdansk Deep, northern SD 26 and SD 28 to the Gotland Basin; inset: location of the study area in a European map.



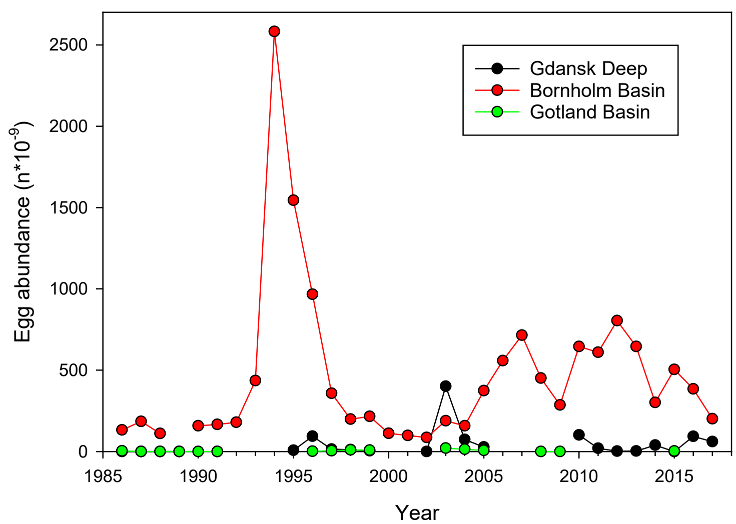


Figure 2. Egg abundances in the Bornholm Basin, Gdansk Deep and Gotland Basin from ichthyoplankton surveys in May/June (upper panel) and July/August (lower panel) 1986–2017.

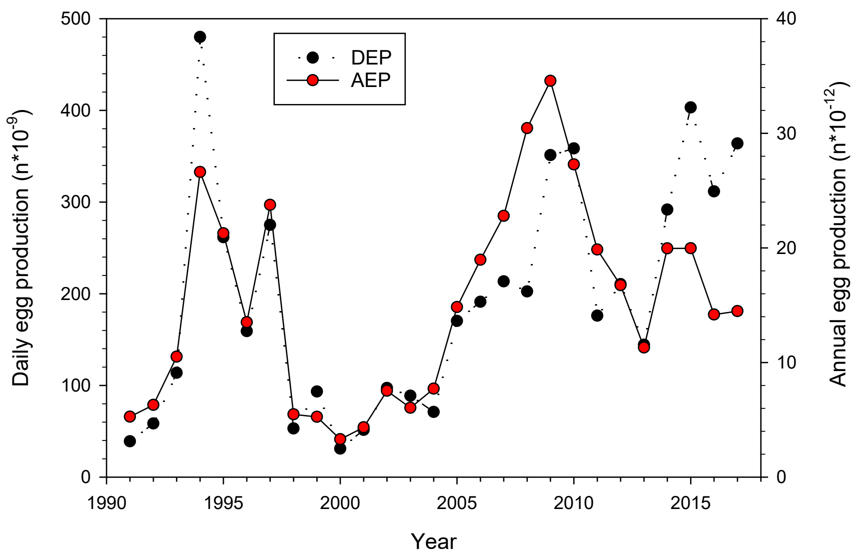


Figure 3. Daily egg production at peak spawning time (DEP) and annual egg production (AEP) in the Bornholm Basin in 1991-2017.

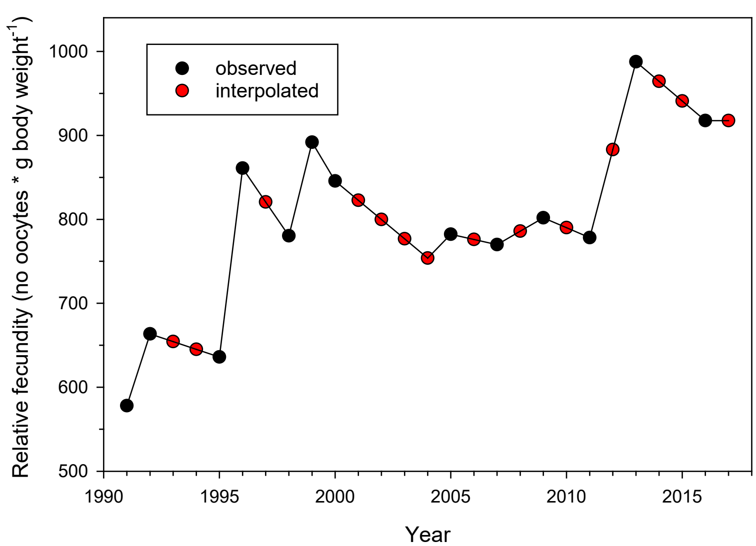


Figure 4. Average relative fecundity observed (black) and interpolated (red) 1991-2017.

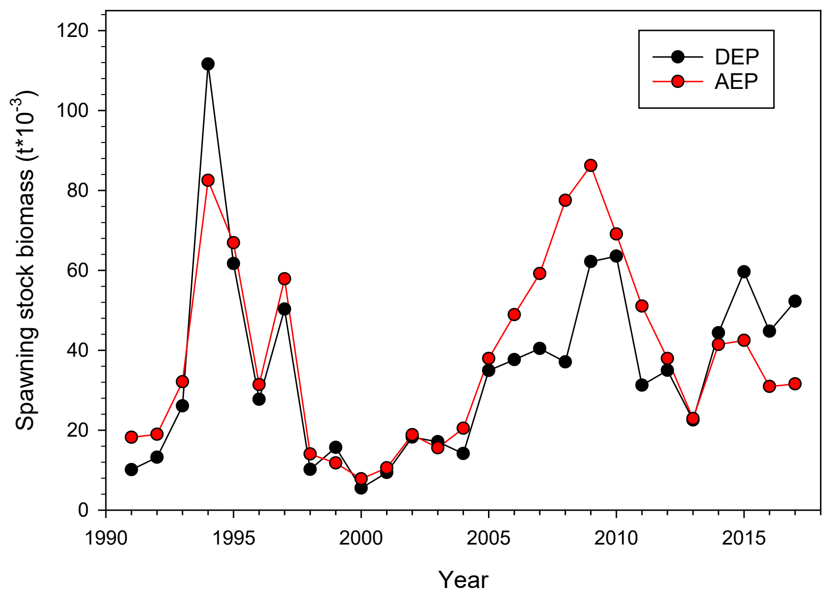


Figure 5. Spawning stock biomass estimated by the Daily egg production (DEP) and the Annual egg production (AEP) method for the Bornholm Basin 1991-2017.

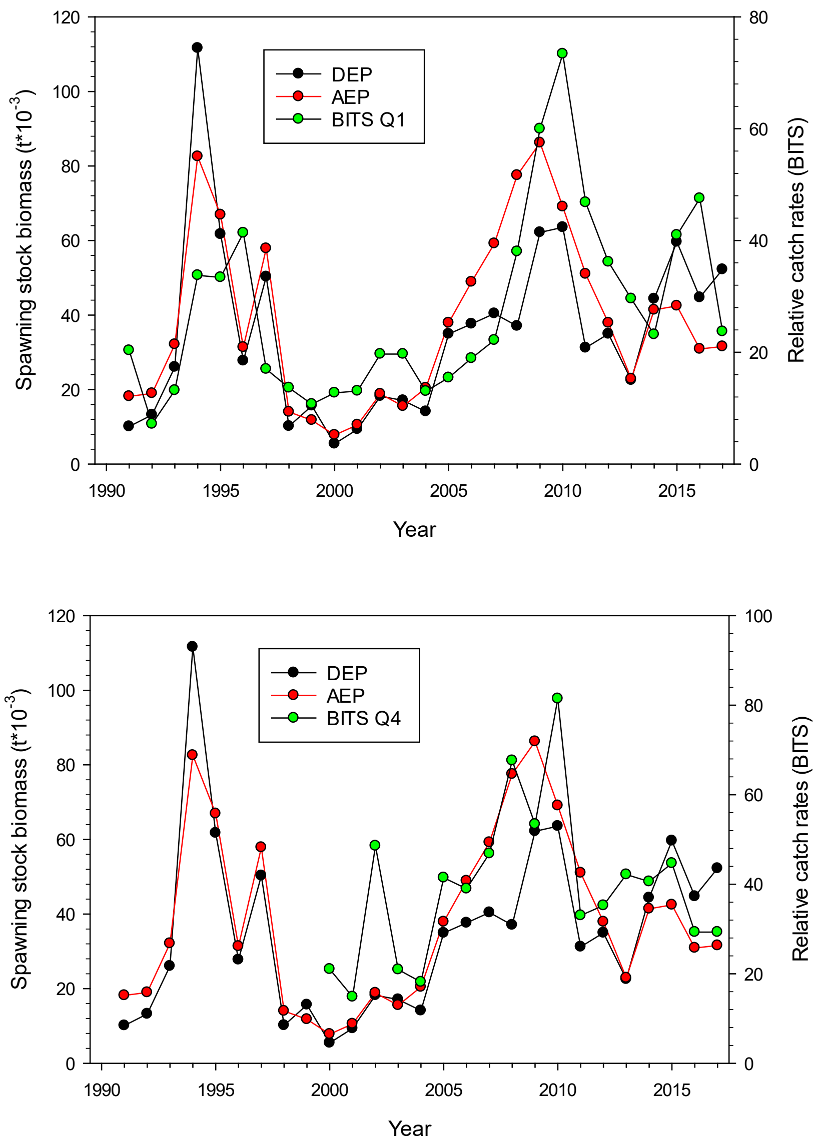


Figure 6. Spawning stock biomass estimated by the Daily egg production (DEP) and the Annual egg production (AEP) method for the Bornholm Basin in comparison to catch rates from the 1. quarter (upper panel) and 4. quarter (lower panel) International Bottom Trawl survey (BITS), the latter as a measure of SSB in the subsequent year; adult cod in the survey were defined as fish ≥30 cm until 2010 and ≥25 cm from 2011 onwards.

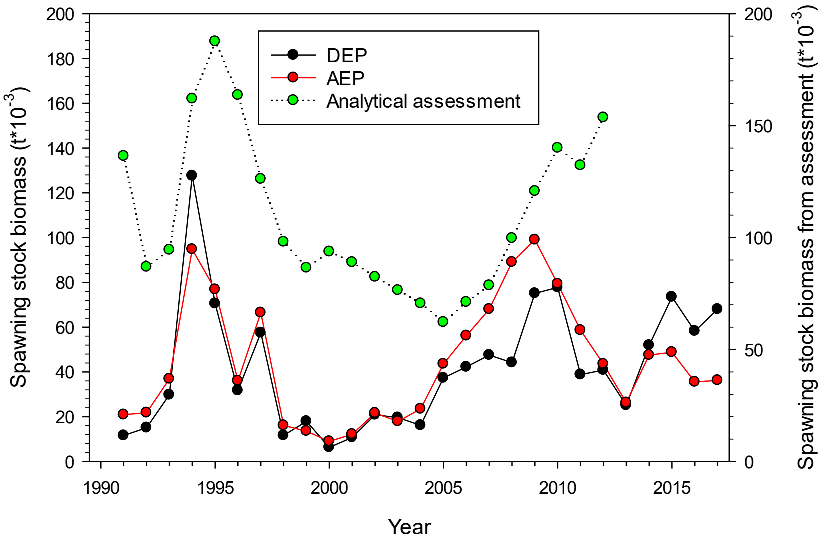


Figure 7. Spawning stock biomass estimated by the Daily egg production (DEP) and the Annual egg production (AEP) method for the Bornholm Basin in comparison to the last accepted analytical assessment (ICES 2014).

WD 4 - Baltic cod larvae index - Description of sampling & index calculation method

Bastian Huwer & Fritz Köster, DTU Aqua, 2018

Sampling, sample preservation and laboratory procedures

Baltic cod larvae are sampled on a fixed station grid in the Bornholm Basin (Fig. 1), consisting of 45 stations (BB-01 to BB-45). The gear in use is a Bongo net (Ø of mouth opening = 60 cm), which is deployed at a speed of 3 kn in a double-oblique haul from the surface to 6 m above the sea floor. The gear is equipped with 1 net of 500 µm mesh size and 1 net of 335 µm mesh size. Both nets are equipped with flowmeters to determine the volume of filtered water in m^3.

The sample from the 335 µm net is preserved immediately after catch in 4% formalin-sea water solution, while the 500 µm sample is usually sorted fresh on board, e.g. to obtain larvae for otolith or biochemical analyses. Only the 335 µm net sample is used for quantitative analyses of larval abundance and calculation of the larvae index. In the laboratory, cod larvae are sorted from the 335 µm samples, counted and measured.

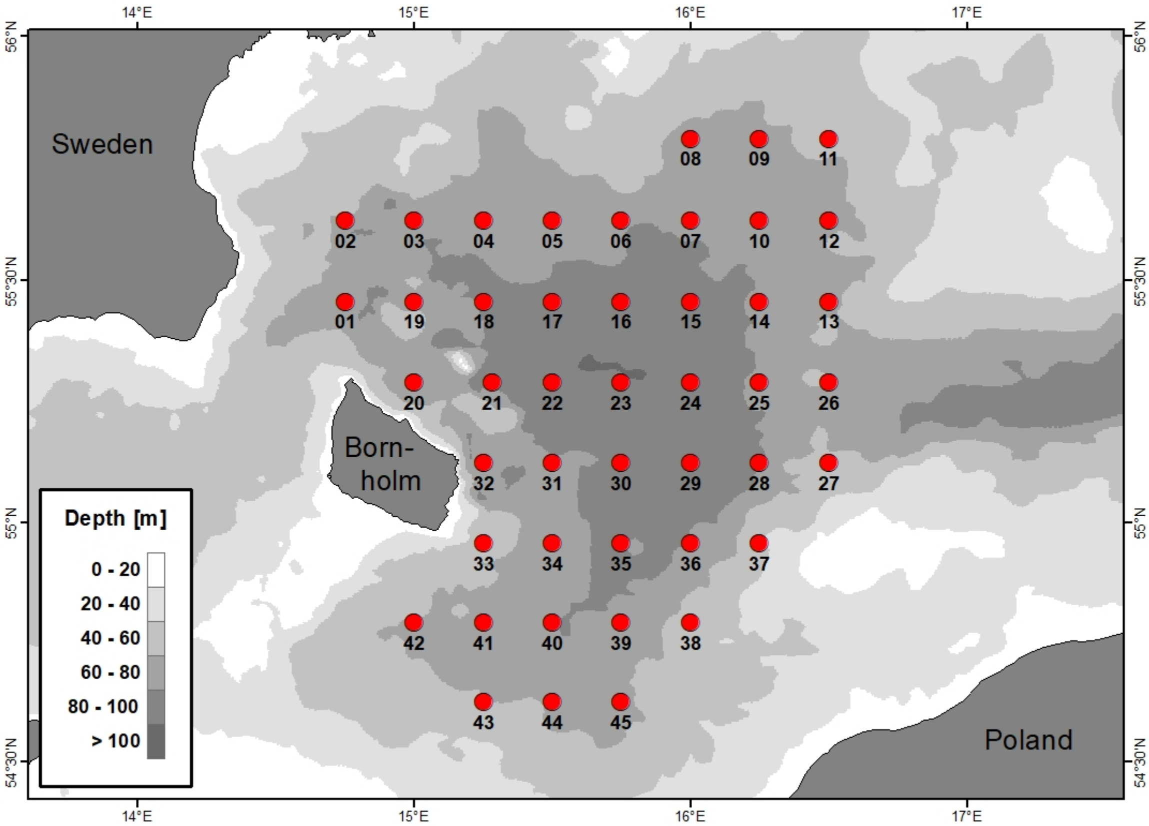


Figure 1: Standard station grid in the Bornholm Basin, consisting of 45 fixed station positions.

Determination of the Baltic cod larvae index (see also example in excel file):

To obtain the larval abundance per m^2 for each station, the number of larvae per haul is first divided by the filtered volume (in m^3) and then multiplied with the gear depth (in m). The average larval abundance per m^2 for each cruise is then determined by averaging the abundance per m^2 from all 45 stations. To obtain the total number of larvae in the survey area, this average abundance per m^2 is then multiplied by the total area of the Bornholm Basin (= 11850000000 m^2). As this results in relatively unhandy numbers for the purpose of an index, the value is finally divided by 1000000000 to obtain the final index (or in other words, the index corresponds to $n*10^9$ larvae).

If data for the other two spawning basins of Eastern Baltic cod (i.e. Gdansk Deep and Gotland Basin) were available, the same procedure was applied to obtain an index for these areas, and the indexes of all 3 basins were summed up. In recent years, the index for the two eastern basins was usually zero or close to zero.

Even though several cruises are conducted each year, the larval index presented and provided for the assessment in connection with WKIDEBCA and WKBALTCOD2 is based on the index value of one cruise at peak spawning time for each year. As peak spawning time has changed from spring to summer over the time series, two different periods of peak spawning time were defined:

For the period 1966-1990 the index is determined for quarter 2 (mainly May/June)

For the period 1991-2017 the index is determined for quarter 3 (mainly July/August)

Some notes concerning potential adjustments/improvements of the index:

As mentioned above, so far the index is based on one cruise at peak spawning per year. However, several cruises are conducted each year, and an index value can be calculated for each of these cruises.

So one could discuss if the index could be improved, e.g. by using the data from all available cruises for each year. Here one could e.g. fit a seasonal larval abundance curve to the data from all cruises and determine the area under this curve (similar as is done for the annual egg production method), and then use the size of this area as index. This may to some degree account for inter-annual variability in spawning time and corresponding time of larval emergence. However, one difficulty could be that cod larvae are generally relatively scarce, i.e. in many years only few or even no larvae are caught on the cruises early and late in the spawning season. This means one would get quite a lot of zero index values for the early and late cruises, which may cause difficulties in the fitting of a seasonal abundance curve.

WD 5 - Historical growth estimates of Eastern Baltic cod from tagging data.

Monica Mion, Alessandro Orio, Roman Motyka, Annelie Hilvarsson, Krzysztof Radtke, Karin Hüssy, Maria Krüger-Johnsen, Kate McQueen, Uwe Krumme, Maris Plikshs and Michele Casini.

Introduction

This document describes the estimation of historical growth parameters of Eastern Baltic cod using a length based method based on tagging data (GROTAG function; Francis, 1988). This method was selected since it incorporates individual variation in growth rate (Tallack, 2009) and for its suitability to handle large tagging datasets, like the historical tagging database for Eastern Baltic cod collated across a number of years. Moreover, the GROTAG function has been successfully applied previously on tagging data to estimate growth rates of cod in the Northeast Atlantic (Tallack, 2009) and western Baltic Sea (McQueen *et al.*, 2018). The first part of the document describes how the historical data have been cleaned to be suitable for growth analyses and the selection of the historical periods, the second part concerns the results of the models to estimate the growth parameters.

Material and methods

Extensive historical external tagging data from tagging experiments performed by Sweden, Germany, Finland, Latvia, Poland and Denmark have been collated. A total of 10278 recaptures (tagged and released fish that have been later recaptured) were available and covered a release period from 1955 to 2006 (Tab. 1; Fig. 1).

Table 1. Overview of historical tagging data available by country.

Country	Release period	Number of recaptures
Denmark	1957-2006	1483
Germany	1962-1981	132
Poland	1957-1970	2299
Sweden	1955-1993	4981
Finland	1974-1984	621
Latvia	1958-1977	762
Total	**1955-2006**	**10278**

The historical tagging activities have been performed mostly in quarter 1, 2 and 4 (Fig. 1)

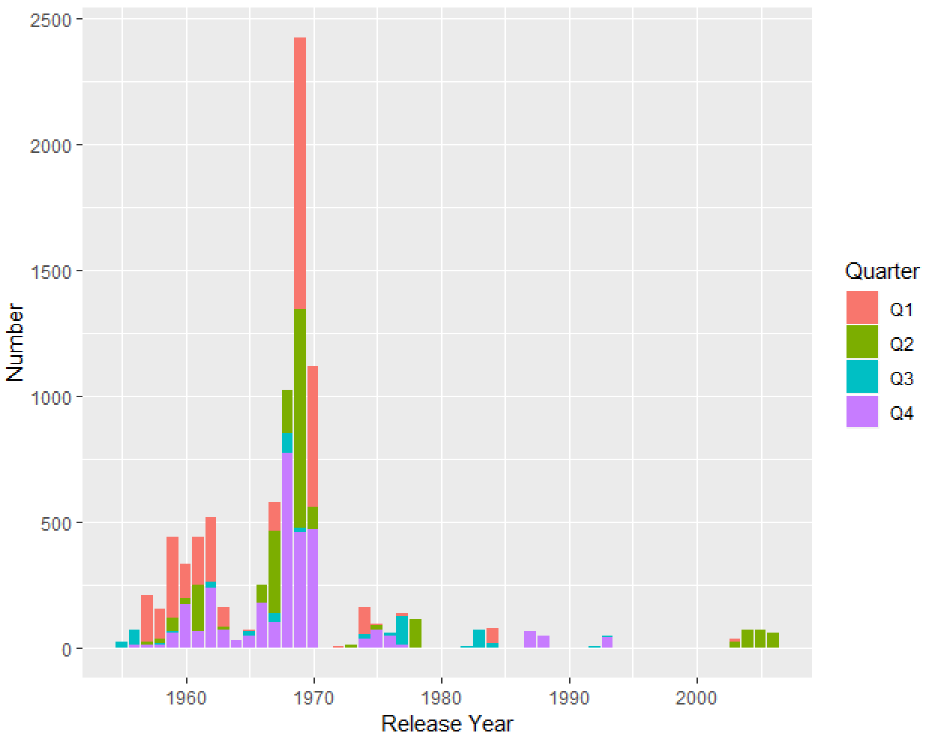


Figure 1. Overview of historical tagging data available by year and quarter.

Data preparation:

Before undertaking growth analyses, some data filters were applied in a stepwise approach. Only fish with tag ID, release and recapture date, location (subdivision) and length measurement were considered. These data were then filtered to omit any cod reported with approximate dates, locations or fish lengths (remaining data n=8856). Fish with negative growth were excluded (remaining data n=7030).

To limit the inclusion of Western Baltic individuals in the Eastern Baltic cod growth analyses, only fish which were both released and recaptured within the boundaries of the Eastern Baltic cod management area (from SD 25 to SD 32) were used (remaining data n=5553). Of these recaptures, to minimize the potential for a downward bias in growth estimate caused by high numbers of short-term recaptures (Fig. 2), only fish with sufficient time between release and recapture (days at liberty, DAL $\geq$ 60) were included in the analyses to ensure enough time for measurable growth to occur (remaining data n=4102). We decided to remove all the recaptures within 60 days from the release following other tagging studies of cod that used as a threshold for measurable growth DAL $\geq$ 60 for Atlantic cod (Tallak, 2009) and DAL $\geq$ 50 for Western Baltic cod (McQueen *et al.,* 2018). Fig. 3 shows that the length distributions of fish with DAL $\geq$ 0 and DAL $\geq$ 60 are similar. A sensitivity analysis was also done to compare the models using the growth estimates with DAL $\geq$ 60 and DAL $\geq$ 90 (Appendix 1).

Figure 2. Overview of different classes of days at liberty (DAL) by year.

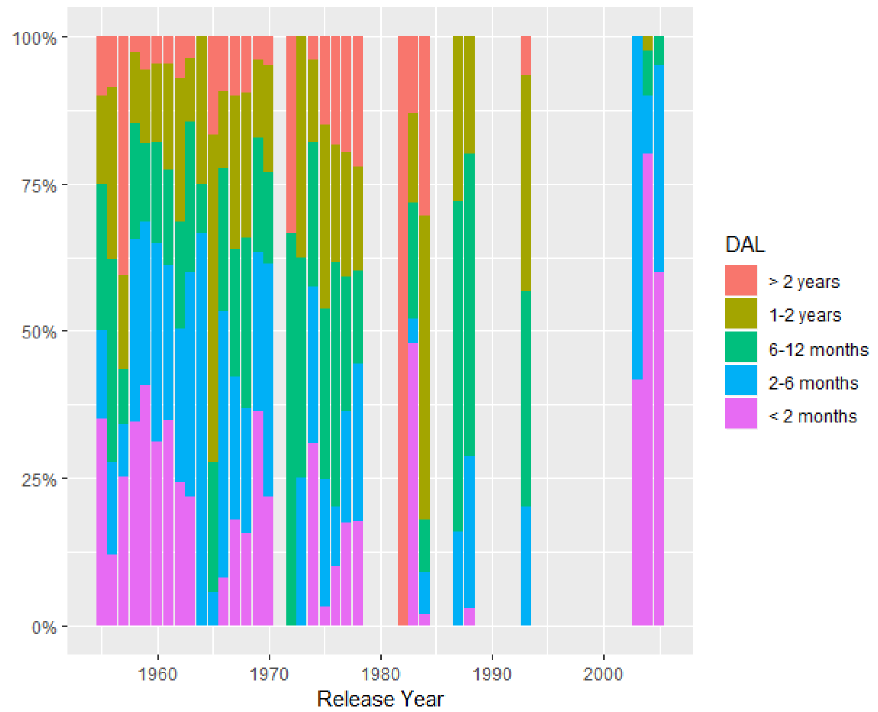


Figure 1. Overview of historical tagging data available by year and quarter.

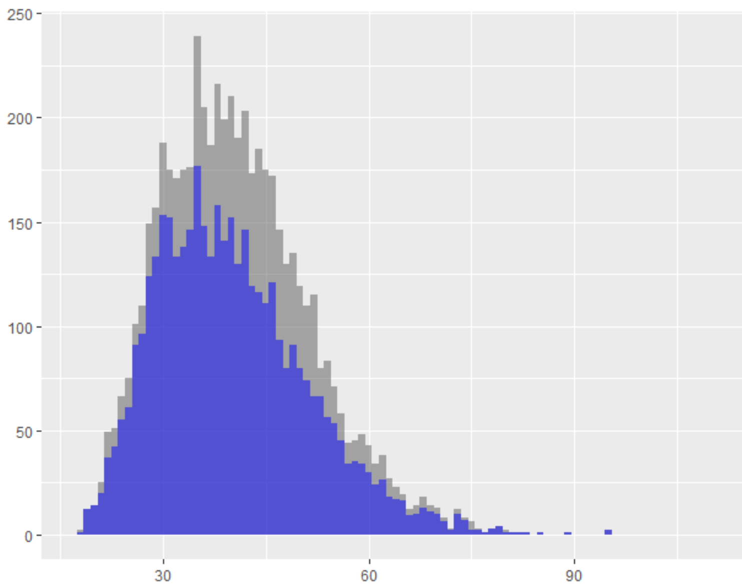


Figure 3. Length distributions at release for recaptured tagged cod. Bars representing cod with DAL ≥ 0 days (n=5553) are shaded grey, and cod with DAL ≥ 60 days (n=4102) are shaded blue.

In addition, we calculated the predicted average annual growth rate (G; Ailloud *et al.,* 2014) of recaptured cod as:

$$G = (\Delta L/\Delta T)*365 \quad (1)$$

where ΔL indicates change in total length of fish and ΔT indicates time-at-liberty in days. The predicted average annual growth rate was then used to exclude extreme growth estimates likely caused by measurement errors. To identify an appropriate maximum annual growth threshold for our data, the Von Bertalanffy (VB) growth parameters for North Sea cod based on Daan *et al.* (1974) were used to

calculate the maximum annual growth for this stock (24.3 cm/y). Thereafter, all the fish in our database with a growth G > 25 cm/y were excluded from further analyses. We decided to take as reference the North Sea cod because of its higher growth than the Eastern Baltic stock. In addition, McQueen *et al.* (2018) estimated G for a 25 cm Western Baltic cod to be around 14 cm/y using GROTAG. Thus, we are confident that, by removing the values with G > 25 cm/y, only the extremes growth estimates likely caused by measurement errors are excluded rather than those caused by individual variability.
After filtering for all these criteria, a total of 3954 tagged cod recaptures were qualified for growth estimation, with releases ranging in size from 13 to 95 cm.

Growth analyses:

Growth parameters were estimated using the GROTAG function in the R library "fishmethods" (Nelsons, 2016) in R v. 3.3.3 (R Core Team, 2017), which is based on Francis (1988). The growth estimation is done using a constrained maximum likelihood optimization. Growth trajectories were compiled from the release length (L1), the recapture length (L2), the time at release (T1), the time at recapture (T2), the change in length between release and recapture (ΔL), and the duration in fraction of years between release and recapture (ΔT). T1 and T2 were measured in fraction of years from 1955, the year of the first tagged cod release.
A total of eight parameters are estimated by the GROTAG function:

1-2) mean annual growth rates (g_α and g_β) in cm/y at two release sizes (α cm and β cm, respectively where α < β). The majority of L1 values must fall within the range of reference lengths α and β (Francis, 1988). Therefore, the same reference lengths (α = 25 cm and β = 55 cm) have been selected within the 5th and 95th percentile values of L1 measurements to easily compare the estimates between time periods.

3) standard deviation of the growth increment (nu).

4) mean measurement error (m).

5) standard deviation of the measurement error (s).

6) contamination probability (p).

7) seasonal growth (w) which describes the time of the year when growth is at its maximum.

8) the magnitude of seasonal growth (u) that ranges from 0 to 1. With u=0 that represents no seasonal growth effect, through u=1 that represents the maximum seasonal growth effect (i.e. where growth ceases over a period each year).

Model selection was done as in Francis (1988), involving incremental combinations of these eight parameters (Tab. 2). The best fitting model (i.e. final model) was selected through Akaike´s Information Criterion (AIC).

Table 2. Parameter combinations estimated by the GROTAG function: five models were applied to the dataset to evaluate optimal model parameterization.

GROTAG model	Estimated parameters
Model 1	g_α, g_β, s
Model 2	g_α, g_β, s, nu
Model 3	g_α, g_β, s, nu, m

Model 4	g_α, g_β, s, nu, m, p
Model 5	g_α, g_β, s, nu, m, p, u, w

Selection of time periods:

Growth parameters were estimated for different time periods (Table 3). The period selection has been made considering the data availability and the main changes in the Eastern Baltic cod stock size based on Eero *et al.* (2007). Thus, four main periods have been selected: 1) 1955-1970 (medium stock size before the cod outburst), 2) 1971-1980 (increase in stock size), 3) 1981-1990 (peak in cod stock size and subsequent decline) and 4) 2003-2006 (after ten years of very low stock size). In addition, growth parameter estimation was also made for 5) 1971-1990, combining the periods 1971-1980 and 1981-1990, and for 6) 1955-1990, combining the periods 1955-1970, 1971-1980 and 1981-1990, i.e. excluding the period of very low cod stock size.

Table 3. Number of data available per country and selected period after the data cleaning.

	Release period					
Country	1955-1970	1971-1980	1981-1990	2003-2006	1971-1990	1955-1990
Germany	1	-	-	-	-	1
Denmark	360	22	44	31	66	426
Poland	860	-	-	-	-	860
Finland	-	342	2		344	344
Latvia	39	68	-		68	107
Sweden	2039	-	117	-	117	2156
Total	**3299**	**432**	**163**	**31**	**595**	**3894**

For the period 1955-1970, length at release was between 17 and 95 cm, length at recapture between 21 and 110 cm (Fig. 4a), and time between release and recapture up to 10.8 years. For the period 1971-1980, length at release was between 24 and 95 cm, length at recapture between 24 and 104 cm (Fig. 4b), and time between release and recapture up to 8.9 years. For the period 1981-1990, length at release was between 21 and 70 cm, length at recapture between 22 and 92 cm (Fig. 4c), and time between release and recapture up to 6.2 years. For the period 1971-1990, length at release was between 21 and 95 cm, length at recapture between 22 and 104 cm (Fig. 4d), and time between release and recapture up to 8.9 years. For the period 1955-1990, length at release was between 17 and 95 cm, length at recapture between 21 and 110 cm (Fig. 4e), and time between release and recapture up to 10.8 years.

After the data cleaning for the period between 2003 and 2006, only 31 recaptures were suitable for growth analyses. Thus, due to the low number of recaptures, and the restricted length range (length at release between 46 and 76 cm, length at recapture between 48 and 77 cm; Fig. 4f), it was not possible to estimate the growth parameters for this time period using the GROTAG model.

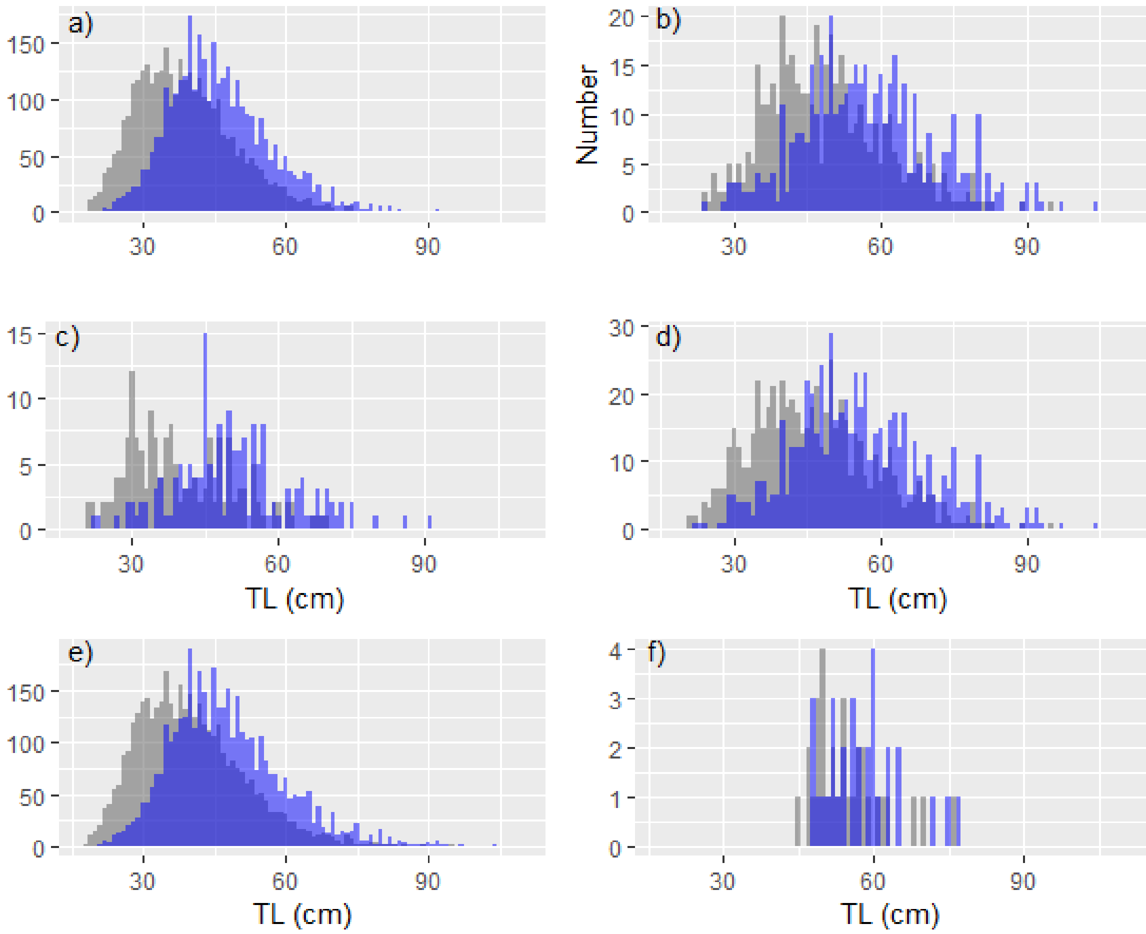


Figure 4. Length distributions at release and recapture of tagged cod for the period (a) 1955-1970, (b) 1971-1980, (c) 1981-1990, (d) 1971-1990, (e) 1955-1990 and (f) 2003-2006. Bars representing release length are shaded grey, and recaptured length is shaded blue

Results

1955-1970:

The full model (i.e. the one including all parameters, model 5) had the lowest AIC value and thus was selected as final model (Tab. 4). The distribution of the model residuals and growth trajectories are presented in Fig. 5. The growth parameters g_α and g_β indicated that growth of cod decreased with increasing length (Tab. 4). The mean growth rates for a small (25 cm) and medium (55 cm) cod were 9.09 cm/y and 6.70 cm/y respectively, as estimated from the growth model parameters. The VB growth parameter estimates derived from the GROTAG function were L_∞=125.27 cm and k =0.10.

Table 4: GROTAG and VB growth parameter estimates (± standard error for the estimated model parameters) calculated for the period 1955-1970 from tagging data for cod α = 25 cm and β = 55 cm (n=3299). Final model is marked bold and "-" indicates whether that the parameter was not included in the model.

		Model 1		Model 2		Model 3		Model 4		**Model 5**	
Parameters		estimate	SE	estimate	SE	Estimate	SE	estimate	SE	**estimate**	**SE**
Mean growth rates	g_α	9.90	0.14	9.11	0.12	9.66	0.17	9.66	0.18	**9.09**	**0.16**
Mean growth rates	g_β	5.18	0.12	6.42	0.10	6.83	0.14	6.84	0.19	**6.70**	**0.33**
Seasonal variation	*u*	-	-	-	-	-	-	-	-	**0.31**	**0.07**
	w	-	-	-	-	-	-	-	-	**0.69**	**0.01**
Growth variability	*nu*	-	-	0.56	0.012	0.53	0.01	0.53	0.02	**0.54**	**0.02**
Measurement error	*s*	4.74	0.06	0.95	0.06	0.92	0.07	0.92	0.07	**1.01**	**0.06**
	m	-	-	-	-	-0.30	0.07	-0.23	0.07	**-0.04**	**0.09**
Outliers	*p*	-	-	-	-	-	-	0.00	0.01	**0.00**	**0.01**
	L_∞	87.94		126.71		127.59		127.60		**125.27**	
	K	0.17		0.09		0.10		0.10		**0.10**	
	AIC	19637.60		17493.50		17476.8		17478.50		**17378.80**	

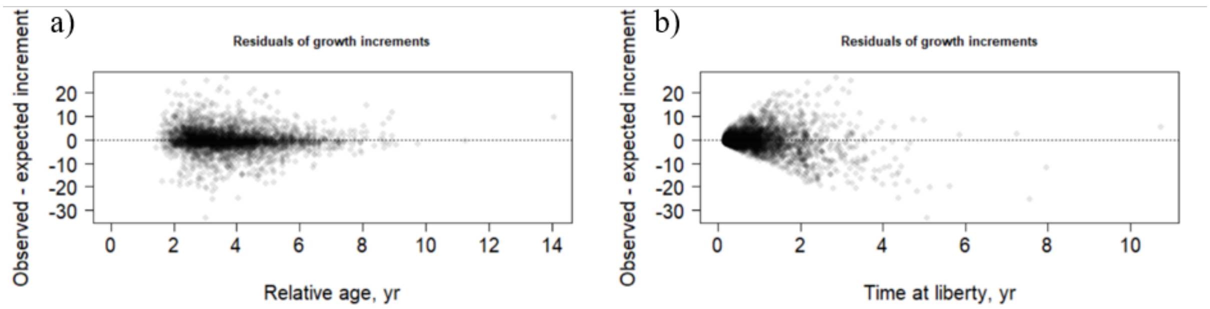


Figure 5: Distribution of the final model residuals (observed-expected growth) plotted against relative age (a) and time at liberty (b).

The growth variability parameter (*nu*) was estimated as 0.54, indicating that individuals within the population could be expected to grow between 0.46 and 1.54 times the estimated average growth (Table 4). The contamination probability (*p*) was negligible (0.00), indicating that the model didn´t detect outliers after the data cleaning. The mean measurement error (*m*) was very low (-0.04 cm) and the standard deviation in measurement error (s) was 1.01 cm, which is in accordance with the 1 cm precision of the length measurements recorded by researchers and fishermen.

1971-1980

For the period 1971-1980, model 4 had the lowest AIC value and thus was selected as final model (Tab. 5). The distribution of the model residuals and growth trajectories are presented in Fig. 5. The growth parameters gα and gβ indicated that growth of cod decreased with increasing length (Tab. 5). The mean growth rates for a small (25 cm) and medium (55 cm) cod were 9.98 cm/y and 6.36 cm/y respectively, as estimated from the growth model parameters. The VB growth parameter estimates derived from the GROTAG function were L∞=107.71 cm and k =0.13.

The growth variability parameter (nu) was estimated as 0.57, indicating that individuals within the population could be expected to grow between 0.43 and 1.57 times the estimated average growth (Table 4). The contamination probability (p) was negligible (0.01), indicating that the model didn´t detect outliers after the data cleaning. The mean measurement error (m) was very low (0.07 cm) and the standard deviation in measurement error (s) was 1.13 cm, which is in accordance with the 1 cm precision of the length measurements recorded by researchers and fishermen.

Table 5: GROTAG and VB growth parameter estimates (± standard error for the estimated model parameters) calculated for the period 1971-1980 from tagging data for cod α = 25 cm and β = 55 cm (n=432). Final model is marked bold and "-" indicates whether that the parameter was not included in the model.

		Model 1		Model 2		Model 3		**Model 4**		Model 5	
Parameters		estimate	SE	estimate	SE	Estimate	SE	**Estimate**	**SE**	estimate	SE
Mean growth rates	g_α	10.01	0.49	10.17	0.46	9.95	0.59	**9.98**	**0.55**	9.38	0.51
Mean growth rates	g_β	5.63	0.17	6.61	0.21	6.45	0.35	**6.36**	**0.34**	6.33	0.34
Seasonal variation	u	-	-	-	-	-	-	**-**	**-**	0.08	0.00
	w	-	-	-	-	-	-	**-**	**-**	0.00	0.00
Growth variability	nu	-	-	0.57	0.03	0.59	0.04	**0.57**	**0.04**	0.57	0.04
Measurement error	s	5.00	0.141	1.34	0.21	1.348	0.21	**1.13**	**0.17**	1.13	0.16
	m	-	-	-	-	0.14	0.24	**0.07**	**0.22**	0.10	0.22
Outliers	p	-	-	-	-	-	-	**0.01**	**0.01**	0.01	0.01
	L_∞	93.49		110.82		110.30		**107.71**		107.02	
	K	0.16		0.13		0.12		**0.13**		0.13	
	AIC	2755.70		2469.60		2471.30		**2462.20**		2465.70	

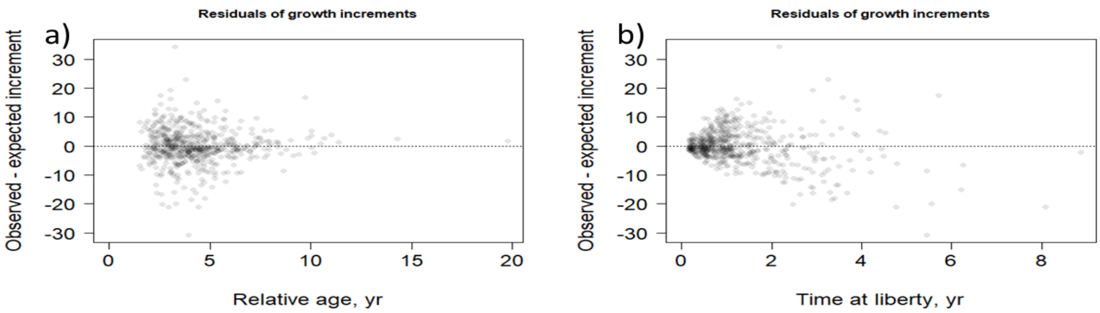


Figure 6: Distribution of the final model residuals (observed-expected growth) plotted against relative age (a) and time at liberty (b) for the period 1971-1980.

1981-1990:

In the time period 1981-1990, model 2 had the lowest AIC and was therefore selected as final model (Tab. 6). However, the limited amount of data prevented the selection of a more complex model that could predict reliable seasonal variability and outliers. The distribution of the model residuals is presented in Fig. 7. The growth parameters g_α and g_β indicate that growth of cod decreased with increasing length (Tab. 6). The mean growth rates for a 25 and 55 cm cod were 13.27 cm/y and 6.50 cm/y, respectively, as estimated from the growth model parameters. VB growth parameter estimates derived from the GROTAG parameters were L_∞=83.66 cm and k=0.26. The growth variability parameter (nu) was estimated as 0.50, indicating that individuals within the population could be expected to grow between 0.50 and 1.50 times the estimated average growth (Tab. 6).

		Model 1		**Model 2**		Model 3		Model 4		Model 5	
Parameters		estimate	SE	**estimate**	**SE**	Estimate	SE	estimate	SE	estimate	SE
Mean growth rates	g_α	13.79	0.68	**13.43**	**0.70**	11.880	1.390	11.760	1.37	10.72	1.25
Mean growth rates	g_β	5.65	0.36	**6.56**	**0.39**	5.740	0.740	5.560	0.74	5.56	0.74
Seasonal variation	u	-	-	**-**	**-**	-	-	-	-	0.00	0.00
	w	-	-	**-**	**-**	-	-	-	-	0.00	0.00
Growth variability	nu	-	-	**0.50**	**0.05**	0.571	0.089	0.528	0.09	0.53	0.09
Measurement error	s	5.00	0.21	**2.111**	**0.58**	1.914	0.686	2.282	0.74	2.29	0.74
	m	-	-	**-**	**-**	1.064	0.814	1.111	0.84	1.121	0.84
Outliers	p	-	-	**-**	**-**	-	-	0.013	0.02	0.013	0.015
	L_∞	75.82		**83.66**		83.04		81.90		81.90	
	K	0.32		**0.26**		0.23		0.23		0.23	
	AIC	1078.40		**1017.30**		1017.80		1018.60		1022.60	

Table 6: GROTAG and VB growth parameter estimates (± standard error for directly estimated model parameters), calculated for the period 1981-1990 from Baltic cod tagging data in SD2532. α = 25 cm and β = 55 cm (n=163). Final model is marked bold and "-" indicates whether the parameter was not included in the model.

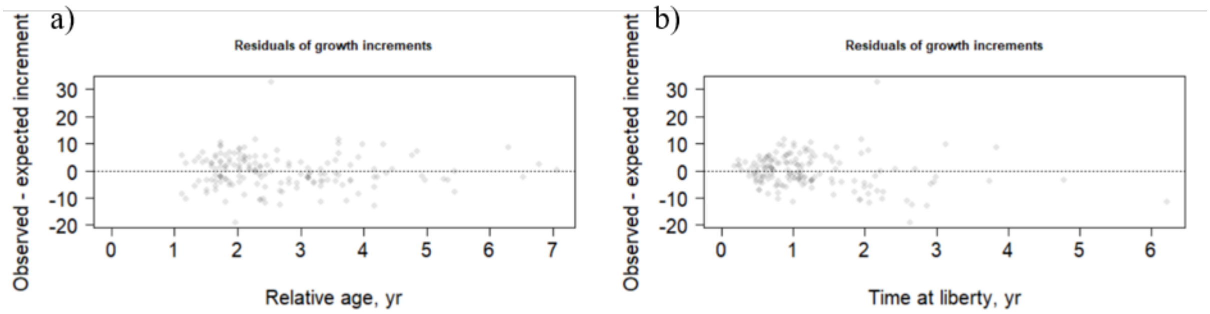


Figure 7: Distribution of the final model residuals (observed-expected growth) plotted against relative age (a) and time at liberty (b) for the period 1981-1990.

1971-1990

In the time period 1971-1990, model 4 had the lowest AIC and was therefore selected as final model (Tab. 7). However, the limited amount of data prevented the selection of a more complex model that could predict reliable seasonal variability and outliers. The distribution of the model residuals is presented in Fig. 8. The growth parameters g_α and g_β indicate that growth of cod decreased with increasing length (Tab. 7). The mean growth rates for a 25 and 55 cm cod were 13.27 cm/y and 6.50 cm/y, respectively, as estimated from the growth model parameters. VB growth parameter estimates derived from the GROTAG parameters were L_∞=83.66 cm and k=0.26. The growth variability parameter (*nu*) was estimated as 0.50, indicating that individuals within the population could be expected to grow between 0.50 and 1.50 times the estimated average growth (Tab. 7).

Table 6**:** GROTAG and VB growth parameter estimates (± standard error for directly estimated model parameters), calculated for the period 1971-1990 from Baltic cod tagging data in SD2532. α = 25 cm and β = 55 cm (n=595). Final model is marked bold and "-" indicates whether the parameter was not included in the model.

		Model 1		Model 2		Model 3		**Model 4**		Model 5	
Parameters		estimate	SE	estimate	SE	estimate	SE	**estimate**	**SE**	estimate	SE
Mean growth rates	g_α	11.35	0.40	11.24	0.36	10.86	0.52	**10.88**	**0.50**	10.02	0.47
Mean growth rates	g_β	5.82	0.15	6.83	0.19	6.56	0.32	**6.49**	**0.32**	6.41	0.32
Seasonal variation	*u*	-	-	-	-	-	-	**-**	**-**	0.18	0.13
	w	-	-	-	-	-	-	**-**	**-**	0.00	0.14
Growth variability	*nu*	-	-	0.55	0.03	0.57	0.03	**0.57**	**0.03**	0.57	0.04
Measurement error	*s*	5.00	0.11	1.51	0.19	1.52	1.92	**1.32**	**0.17**	1.32	0.19
	m	-	-	-	-	0.246	0.242	**0.19**	**0.02**	0.24	0.23
Outliers	*p*	-	-	-	-	-	-	**0.01**	**0.01**	0.01	0.01
	L_∞	86.58		101.39		100.85		**99.38**		99.32	
	k	0.20		0.16		0.15		**0.16**		0.15	
	AIC	3848.90		3498.80		3499.70		**3493.90**		3494.60	

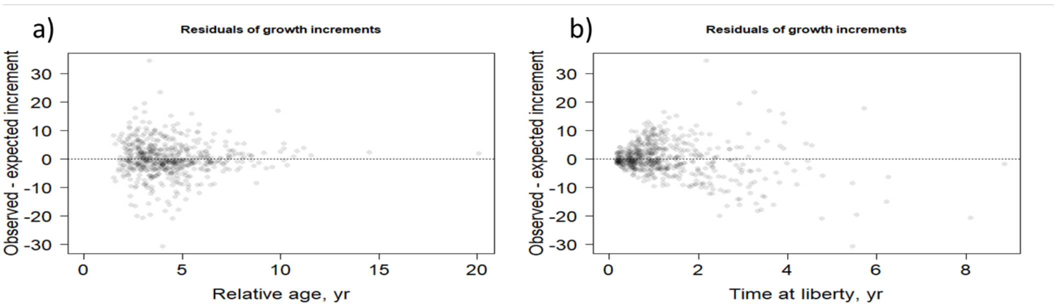


Figure 8: Distribution of the final model residuals (observed-expected growth) plotted against relative age (a) and time at liberty (b) for the period 1971-1990.

1955-1990:

The full model (i.e. model 5) had the lowest AIC value and thus was selected as final model (Tab. 8). The distribution of the model residuals are presented in Fig. 9. The growth parameters g_α and g_β indicate that growth of cod decreases with increasing length (Table 8). The mean growth rates for a 25 and 55 cm cod are 9.57 cm/y and 6.53 cm/y, respectively, as estimated from the growth model parameters. VB growth parameter estimates derived from the GROTAG function were L_∞=108.56 cm and k =0.13. The growth variability parameter (*nu*) was estimated as 0.53, indicating that individuals within the population could be expected to grow between 0.47 and 1.53 times the estimated average growth per length class (Table 8). The contamination probability (*p*) was negligible (0.00), indicating that the model didn´t detect outliers after the data cleaning. The mean measurement error (*m*) was very low (-0.09 cm) and the standard deviation in measurement error (s) was 1.04 cm that is in accordance with the 1 cm precision of the length measurements recorded by researchers and fishermen. However, the estimates for the whole period 1955-1990 are more similar to estimates for the period 1955-1970 because of the higher amount of data belonging to the first period (1955-1970) compared to the following periods (1971-1990; Tab. 8).

Table 8: GROTAG and VB growth parameter estimates (± standard error for directly estimated model parameters), calculated for the period 1955-1990 from Baltic cod tagging data in SD2532. α = 25 cm and β = 55 cm (n=3443). Final model is marked bold and “-” indicates whether the parameter was not included in the model.

		Model 1		Model 2		Model 3		Model 4		**Model 5**	
Parameters		estimate	SE	estimate	SE	estimate	SE	estimate	SE	**estimate**	**SE**
Mean growth rates	g_α	10.13	0.14	9.32	0.12	9.91	0.17	9.91	0.17	**9.57**	**0.16**
Mean growth rates	g_β	5.23	0.11	6.47	0.10	6.91	0.13	6.91	0.14	**6.53**	**0.16**
Seasonal variation	*u*	-	-	-	-	-	-	-	-	**0.30**	**0.03**
	w	-	-	-	-	-	-	-	-	**0.70**	**0.01**
Growth variability	*nu*	-	-	0.57	0.01	0.53	0.01	0.53	0.01	**0.53**	**0.01**
Measurement error	*s*	4.81	0.06	0.96	0.06	0.92	0.07	0.92	0.07	**1.04**	**0.06**
	m	-	-	-	-	-0.32	0.07	-0.32	0.07	**-0.09**	**0.07**
Outliers	*p*	-	-	-	-	-	-	-	-	**0.00**	**0.00**
	L_∞	87.02		123.21		124.01		124.01		**108.56**	
	k	0.18		0.10		0.11		0.11		**0.13**	
	AIC	20587.10		18434.70		18415.20		18417.20		**18329.10**	

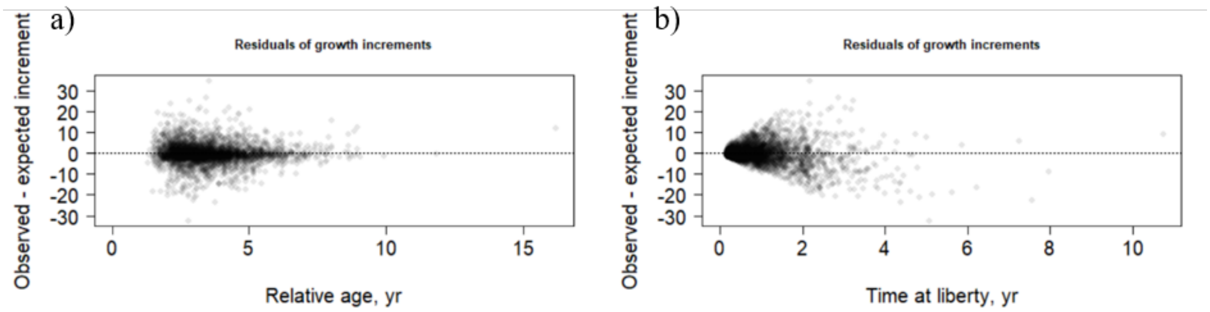


Figure 9: Distribution of the final model residuals (observed-expected growth) plotted against relative age (a) and time at liberty (b) for the period 1955-1990.

Growth analyses by area:

Possible differences in growth related to different areas were investigated: north (SDs 29-32) EB and central-southern EB (SD 25-28). The period 1955-1975 was selected to be able to have enough data in both areas. For the north (SDs 29-32) EB L_∞ was 97.65 cm (n=215), for the central-southern EB (SD 25-28) L_∞ was 126.20 cm (n=3267). Other combinations of subdivisions have been tried and the general trend is a smaller L_∞ for the northern area compare to the southern area.
In general, for all the different periods analysed, north and south are covered but the proportion of samples in the areas are unbalanced and this might drive the estimates towards smaller or larger L_∞ depending on the proportion of recaptures in the different areas.

Sensitivity analysis:

A sensitivity analysis was also done comparing the measurement error between the final model with DAL ≥ 60 and DAL ≥ 90 days for the period 1955-1970, 1971-1980 and 1981-1990. The measurement error was lower for the models with DAL ≥ 60 for the period 1955-1970, 1971-1980 but not for the period 1981-1990 (this is probably due to the lower number of recaptures available for the analysis; Appendix 1).

Acknowledgements:

We thanks the Swedish Agency for Marine and Water Management for financing the compilation of the data. Funds for data compilation and analyses were also provided by the TABACOD project financed by the BalticSea2020 foundation (http://balticsea2020.org/)".

References:

Daan, N., 1974. Growth of North Sea Cod, *Gadus morhua*. Netherlands Journal of Sea Research 8(1):27-48.

Eero, M., Köster, F. W., Plikshs, M., and Thurow, F., 2007. Eastern Baltic cod (*Gadus morhua callarias*) stock dynamics: extending the analytical assessment back to the mid-1940s, ICES Journal of Marine Science, Volume 64(6):1257–1271.

Francis, R.I.C.C. 1988. Maximum likelihood estimation of growth and growth variability from tagging data. New Zealand journal of marine and freshwater research, 22(1), 43-51.

McQueen, K., Eveson, J. P., Dolk, B., Lorenz, T., Mohr, T., Schade, F. M., Krumme, U., 2018. Growth of cod (*Gadus morhua*) in the western Baltic Sea: estimating improved growth parameters from tag recapture data. Canadian Journal of Fisheries and Aquatic Sciences. https://doi.org/10.1139/cjfas-2018-0081

Nelson, G.A., 2016. fishmethods: fishery science methods and models in R. R package version 1.10-626 0. https://CRAN.R-project.org/package=fishmethods

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Tallack, S.M.L., 2009. Regional growth estimates of Atlantic cod, *Gadus morhua*: Applications of the maximum likelihood GROTAG model to tagging data in the Gulf of Maine (USA/Canada) region. Fisheries Research, 99:137-150.

Appendix:

Appendix 1. GROTAG and VB growth parameter estimates (± standard error for the estimated final model parameters) calculated for the period 1955-1970, 1971-1980 and 1981-1990 from Baltic cod tagging data in SD2532 with DAL ≥ 60 and DAL ≥ 90, α = 25 cm and β = 55 cm.

		1955-1970				1971-1980				1981-1990			
		n=3299		n=2909		n=432		n=394		n=163		n=160	
		DAL ≥ 60		DAL ≥ 90		DAL ≥ 60		DAL ≥ 90		DAL ≥ 60		DAL ≥ 90	
		Model 5		Model 5		Model 4		Model 4		Model 2		Model 2	
Parameters		estimate	SE	estimate	SE	estimate	SE	estimate	SE	estimate	SE	estimate	SE
Mean growth rates	*gα*	9.090	0.160	9.310	0.190	9.980	0.550	9.760	0.620	13.430	0.700	13.330	0.700
Mean growth rates	*gβ*	6.700	0.330	6.320	0.170	6.360	0.340	6.090	0.390	6.560	0.390	6.540	0.390
Seasonal variation	*u*	0.306	0.065	0.346	0.039	-	-	-	-	-	-	-	-
	w	0.689	0.013	0.668	0.016	-	-	-	-	-	-	-	-
Growth variability	*nu*	0.540	0.017	0.507	0.014	0.568	0.039	0.586	0.047	0.496	0.048	0.511	0.055
Measurement error	*s*	1.014	0.061	1.383	0.077	1.129	0.165	1.200	0.250	2.111	0.582	1.834	0.844
	m	-0.042	0.091	-0.025	0.111	0.070	0.220	0.356	0.308	-	-	-	-
Outliers	*p*	0.000	0.014	0.002	0.001	0.014	0.009	0.017	0.010	-	-	-	-
	$L\infty$	-125.274		-107.911		107.715		104.820		83.664		83.918	
	k	0.100		0.127		0.129		0.130		0.260		0.256	

WD 6 - Estimation of natural mortality for Eastern Baltic cod

By Massimiliano Cardinale

Department of Aquatic Resources (SLU Aqua), Institute of Marine Research

Introduction

This working document describes the analysis carried out to estimate the natural mortality by age for Eastern Baltic cod as derived using information on longevity, growth, maturity and other life trait history parameters. In the last accepted assessment for Eastern Baltic cod (Cod (*Gadus morhua*) in subdivisions 24–32, eastern Baltic stock (eastern Baltic Sea)) carried out in 2013 the stock was assessed using a SAM model (Berg and Nielsen 2016), with a value of natural mortality (M) equal to 0.20 for all ages and years. However, the value of M equal to 0.20, which has been used since 1988, was not derived from longevity, growth, maturity or other life trait history parameters but it was based on the analysis made by Grzebielec and Kosior (1987), where M was estimated between 0.12-0.18.

Before 1988, the ICES Baltic Assessment Working Group was not able to derive a more precise estimate of M, which was assumed to be 0.3. This value was also not derived from longevity, growth, maturity or other life trait history parameters, but it was based on stomach samples and cannibalism (Anon., 1988b). The 1987 Baltic Multispecies Assessment working group (Anon., 1987b) using stomach data for the period 1982-1984, disregarded the predation on cod as very few cod were found in the stomachs during that period and thus M = 0.3 was used. However, in 1988 (Anon., 1988a and b) a reduced M = 0.2 was assumed, which was based on the rationale that the amount of food available for cod has not changed and thus cannibalism should not be expected. The best documentation of the historical choice of M for the assessment of Eastern Baltic cod and the reason for changing from 0.3 to 0.2 is available in Anon., (1988a).

Materials and methods

Different methods were applied to estimate natural mortality for Eastern Baltic cod. The Hoenig method (1983) was applied to derive M for Eastern Baltic cod and it is based only on maximum age for teleosts. The maximum observed age (*tmax*) for Eastern Baltic cod as recorded during BITS survey since 1991 is 15 years. This is derived from more than 172 536 age readings from all countries participating in the survey. Moreover, data from fish caught by the commercial fisheries over the last decades showed a maximum age of 20 years. Therefore, a maximum age of 20 years was used to estimate natural mortality for Eastern Baltic cod. The Hoenig method (1983) gave an M of 0.223 for a maximum age of 20 years for teleosts (Table 1). A more recent paper by Then et al., (2015) analysed data from 226 studies (including Hoenig 1983) to evaluate the robustness of life-history based M inferences. Based on updating and testing indirect estimators of natural mortality using information on 201 fish species, Then et al., (2015) recommend the use of their updated maximum age-based estimator when possible and an updated von Bertalanffy K-based method otherwise. Despite concerns remaining over these methods, Then et al., (2015) advice is probably the best currently available, except in cases for which reliable direct estimates of M are available or simulation analyses indicate that estimates within the stock assessment model are reliable.

Results and discussion

Then et al., (2015) *tmax* based method (i.e. $M = 4.899 \cdot tmax^{-0.916}$) gives an M value of 0.315 for a *tmax* of 20 years (Table 1). Then et al., (2015) von Bertalanffy K-based method, which uses the parameters of the von Bertalanffy growth curve, K and Linf ($M = 4.118.K^{0.73}.Linf^{-0.33}$), predicts M = 0.156 for Eastern Baltic cod. The growth parameters (k=0.10 and L_{inf} = 125.27 cm) were estimated from historical tagging data (See Mion et al., working document in this report for details). Here we propose to use both methods as suggested by Then et al., (2015), which are based on maximum age (*tmax*) and parameters of the Bertalanffy growth curve. However, for completeness, inferences on natural mortality estimates from a range of life-history based methods are presented in Table 1. The history traits parameters used in the M calculations are shown in Table 2.

Natural mortality is high in young fish and declines with age, as shown by multispecies models that include diet data and estimation of size preferences such as applied by ICES WGSAM for the North Sea and the Baltic Sea (ICES, 2018). Proxy methods to infer age-dependent M in younger fish are given by Lorenzen (1996) and Gislason et al., (2010). The Gislason method generally gives lower M for adult fish. Brodziak et al (2011) suggest that methods such as Lorenzen can be used to derive the relative age-dependent patterns for younger fish, but can be re-scaled to give M at older ages that are more similar to those from methods using (e.g.) *tmax*. Therefore, Lorenzen (1996) method was used to estimate age-dependent M values for Eastern Baltic cod and the results are given in Fig. 1 and Tables 2. Fig. 1 and Table 3 show Lorenzen M values rescaled to give mean M at ages 10-15, which are equivalent to the Then et al., (2015) prediction of 0.315 and 0.156 for *tmax* 20 years old and growth based method, respectively. Therefore, for the benchmark, the following M options could be explored:

1. M = 0.156 at all ages (Then et al., 2015)
2. M = 0.315 at all ages (Then et al., 2015)
3. Lorenzen M (age specific) rescaled to M=0.156
4. Lorenzen M (age specific) rescaled to M=0.315
5. M= 0.2 at all ages for continuity with previous stock assessments.

Tables and figures

Table 1. Eastern Baltic cod (Cod (*Gadus morhua*) in subdivisions 24–32, eastern Baltic stock (eastern Baltic Sea)). Inferences on natural mortality rate from a range of life-history traits based methods. The growth parameters used for the calculations (k=0.10 and L_{inf} = 125.27 cm) were estimated from historical tagging data (See Mion et al., working document).

Source	Formulation	Combined sex M		
				tmax
Hoenig 1983	variety of taxa ln(M) = 1.44-0.982*ln(tmax);			0.223
	teleosts ln(M) = 1.46-1.01*ln(tmax)			0.209
	M= 4.899*tmax^-.916 (from 226 species)			0.315
Then *et al* 2015	M= 4.118*K^0.73. Linf^-0.33	0.156		
Alverson and Carney 1975	M = 3k/(exp(0.38*tmax*k)-1)			0.264
Pauly 1980	M=exp(-0.0152+0.6543*ln(k)-0.279*ln(Linf,cm)+0.4634*lnT(oC))	0.130	TdegC=	6
		0.140	TdegC=	7
		0.149	TdegC=	8
Ralston 1987	M=0.0189+2.06*k	0.225		
Beverton 1992	M=3k/(exp(*am* *k)-1) am = age at 50% maturity	0.857	female *am* ; comb sex k	
		1.355	male *am* , comb sex k	
Jensen (1997)	M=1.5K	0.150		
			Gislason	Lorenzen
Gislason 2010	M = exp(0.55-1.61*Ln(L) + 1.44* Ln (Linf)+ Ln(K))	age 1	1.804	0.959
Lorenzen	M=3*W^-0.288	age 2	0.860	0.488
		age 5	0.304	0.353
	Gislason: L = length at age from VBGF	age 7	0.213	0.289
	Lorenzen: W = mean wt at age from length-weight relationship	age 10	0.152	0.253
		age 15	0.112	0.202
		age 20	0.095	0.184

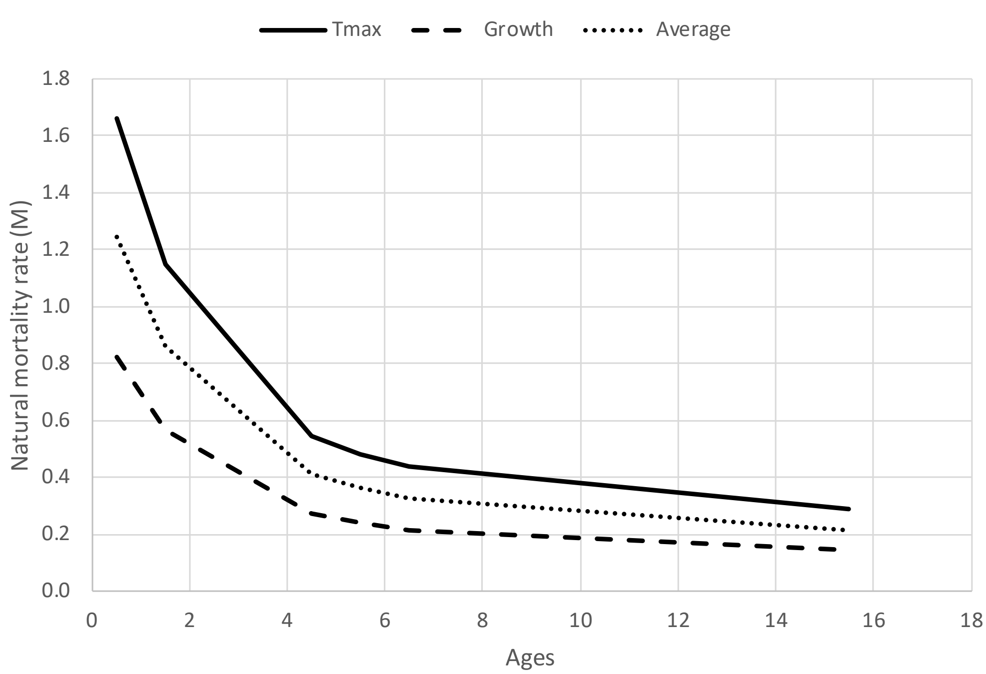
Table 2. Eastern Baltic cod (Cod (*Gadus morhua*) in subdivisions 24–32, eastern Baltic stock (eastern Baltic Sea)). Inferences on natural mortality rate from a range of life-history traits based methods. The growth parameters used for the calculations (k=0.10 and L_{inf} = 125.27 cm) were estimated from historical tagging data (See Mion et al., working document).

Life history parameters	
k (combined sex)	0.1
L_{inf} (combined sex)	125.27
t_o (combined sex)	-0.51
Age at 50% maturity females	3
Age at 50% maturity males	2
Max age (combined sex)	20
Length at 50% mat (females)	35
Length at 50% mat (males)	25
a	6.58E-06
b	*3.1353*

Table 2. Eastern Baltic cod (Cod (*Gadus morhua*) in subdivisions 24–32, eastern Baltic stock (eastern Baltic Sea)): Natural mortality by age values estimated using: Lorenzen (1996) (Lorenzen); Then et al., 2015 maximum age method (*tmax*) rescaled to a mean M of 0.315 at ages 10-15 based on a maximum age of 20 years; Then et al., 2015 growth method rescaled to a mean M of 0.156 at ages 10-15 based on a maximum age of 20 years. The growth parameters used for the calculations (k=0.10 and L_{inf} = 125.27 cm) were estimated from historical tagging data (See Mion et al., working document).

Age	*Tmax*	Growth	Average
0.5	1.663	0.822	1.243
1.5	1.147	0.567	0.857
4.5	0.547	0.270	0.409
5.5	0.483	0.239	0.361
6.5	0.437	0.216	0.326
15.5	0.287	0.142	0.215

Fig. 1. Eastern Baltic cod (Cod (*Gadus morhua*) in subdivisions 24–32, eastern Baltic stock (eastern Baltic Sea)): Natural mortality by age values inferred from Lorenzen (1996) rescaled to a mean M of 0.315 at ages 10-15 (based on Then et al 2015 maximum age method, for a maximum age of 20 years) and mean M of 0.156 at ages 10-15 (based on Then et al., 2015 growth method). The growth parameters used for the calculations (k=0.10 and L_{inf} = 125.27 cm) were estimated from historical tagging data (See Mion et al., working document).



References

Anon. 1987a. Report of the Working Group on Demersal Stocks in the Baltic. ICES, Assess:22. Assessment of Doc. C.M.1987.

Anon. 1987b. Report of the Baltic Multispecies Assessment Working Group. lCES, Doc. C.M.1987/Assess:6.

Anon. 1988a. Report of the Working Group on Demersal Stocks in the Baltic. ICES, Assess:20. Assessment of Doc. C.M.1988.

Anon. 1988b. Report of the Working Group on Multispecies Assessments of Baltic Fish. ICES, Doc. C.M.1988/Assess:4.

Brodziak, J., Ianelli, J., Lorenzen, K., Methot, R. D. 2011. Estimating natural mortality in stock assessment applications. NOAA Technical Memorandum NMFS-F/SPO-119pg. 38 pp.

Grzebielec. R., and Kosior, M. 1987. Mortality of cod in the Southern Baltic in 1973-1986. ICES CM 1987/J: 5.

Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology, 49: 627–642.

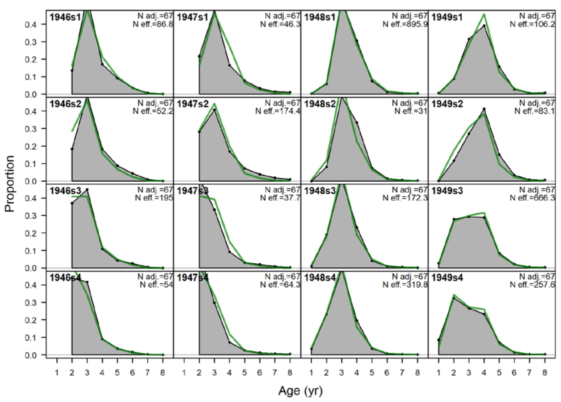
Nielsen, A., and Berg, C. W. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research, 158: 96-101.

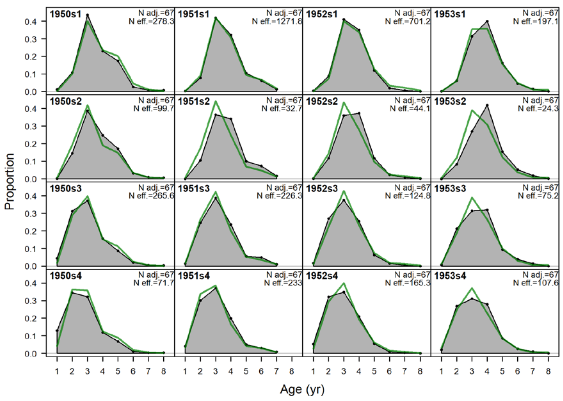
158: 96 –101. Then, A.Y., J.M. Hoenig, N.G. Hall and D.A. Hewitt. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES Journal of Marine Science 72 (1): 82-92. doi: 10.1093/icesjms/fsu136.

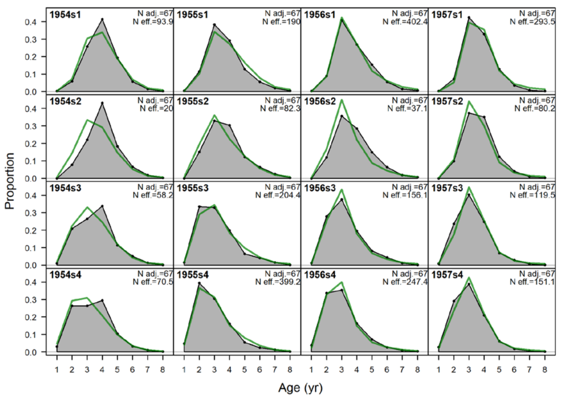
ICES. 2018. Interim Report of the Working Group on Multispecies Assessment Methods (WGSAM), 16–20 October 2017, San Sebastian, Spain. ICES CM 2017/SSGEPI:20. 395 pp.

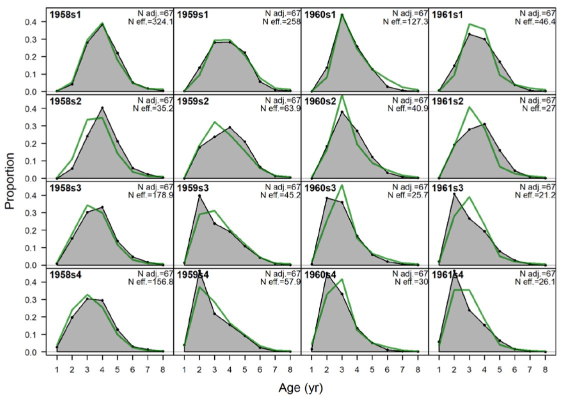
WD 7 - Fits to age, length and conditional age at length (ALK) data for final Stock Synthesis run.

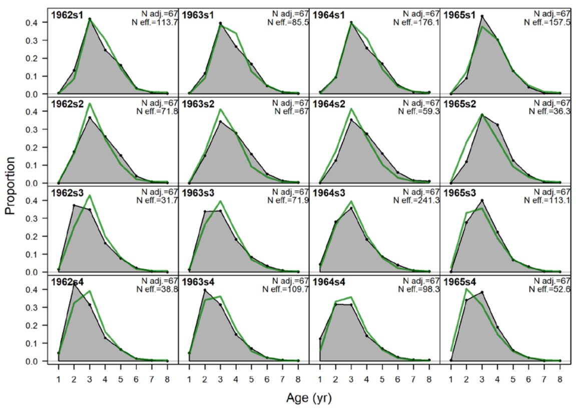
Fits to age compositions: Active, Passive

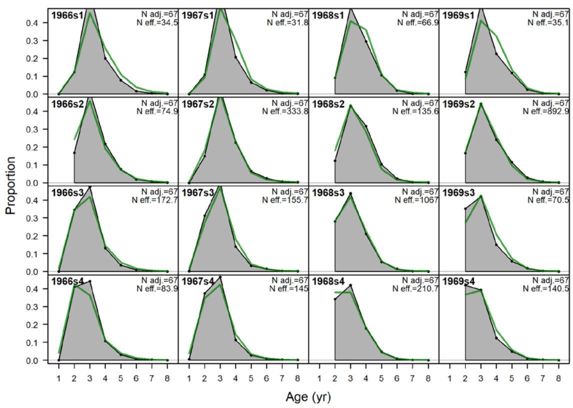


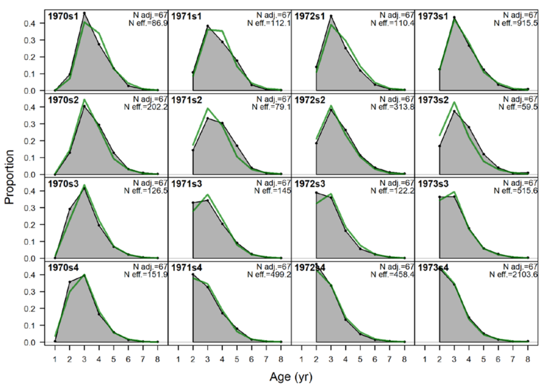


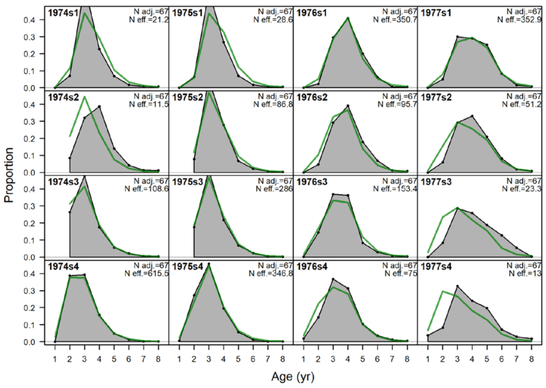


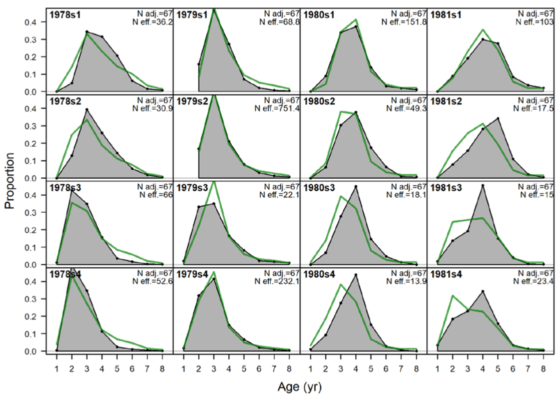


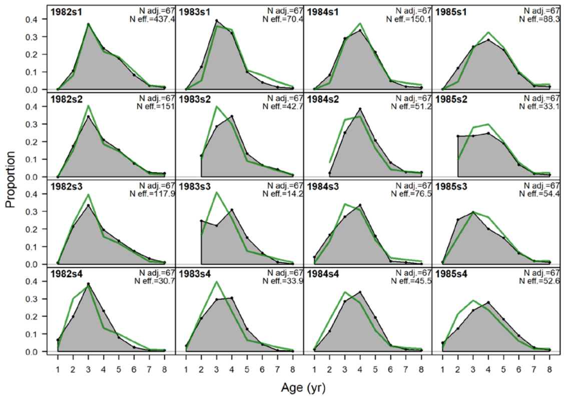


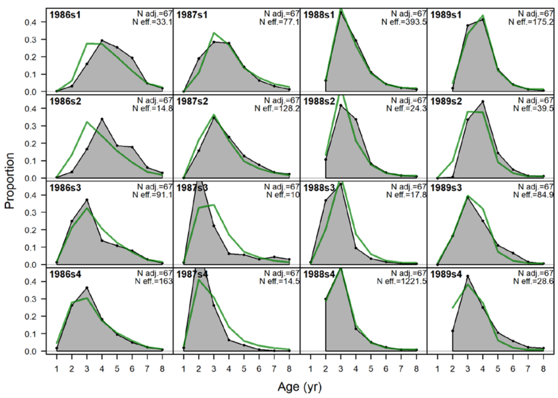


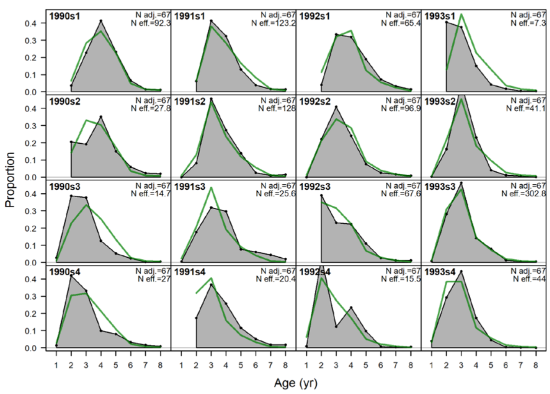

1970s1
1971s1
1972s1
1973s1
1970s2
1971s2
1972s2
1973s2
1970s3
1971s3
1972s3
1973s3
1970s4
1971s4
1972s4
1973s4
Proportion
Age (yr)

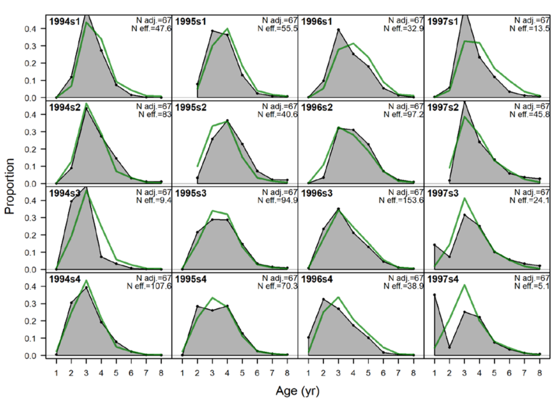

1974s1
1975s1
1976s1
1977s1
1974s2
1975s2
1976s2
1977s2
1974s3
1975s3
1976s3
1977s3
1974s4
1975s4
1976s4
1977s4
Proportion
Age (yr)

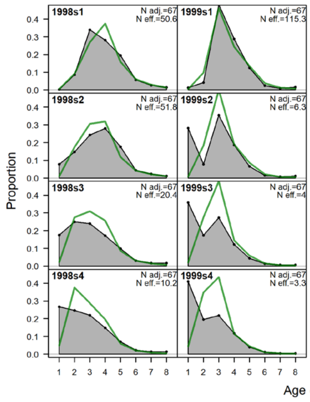

1978s1
1979s1
1980s1
1981s1
1978s2
1979s2
1980s2
1981s2
1978s3
1979s3
1980s3
1981s3
1978s4
1979s4
1980s4
1981s4
Proportion
Age (yr)

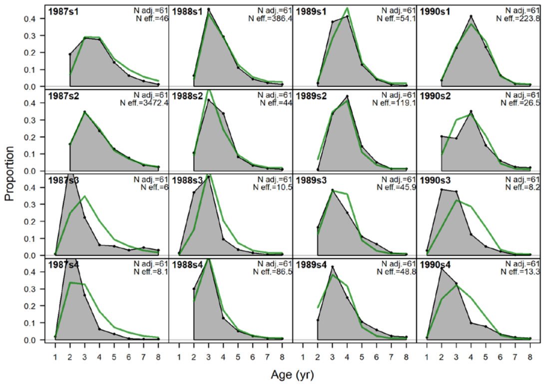

1982s1
1983s1
1984s1
1985s1
1982s2
1983s2
1984s2
1985s2
1982s3
1983s3
1984s3
1985s3
1982s4
1983s4
1984s4
1985s4
Proportion
Age (yr)

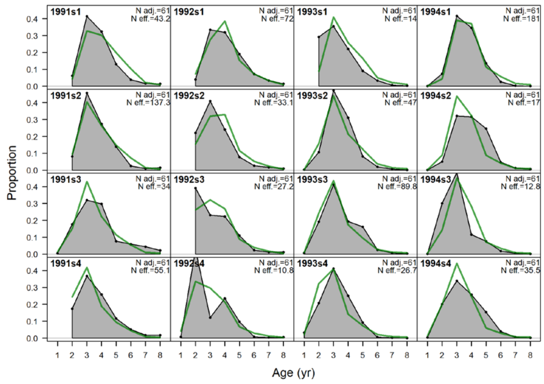

1986s1
1987s1
1988s1
1989s1
1986s2
1987s2
1988s2
1989s2
1986s3
1987s3
1988s3
1989s3
1986s4
1987s4
1988s4
1989s4
Proportion
Age (yr)

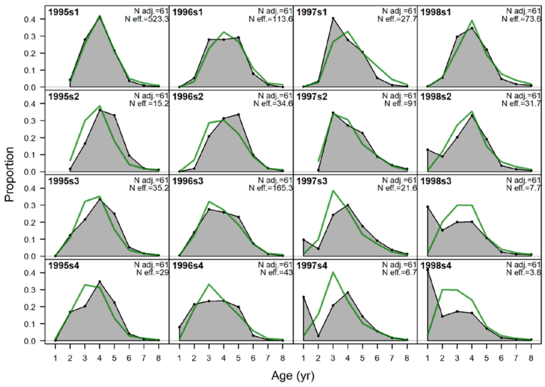

1990s1
1991s1
1992s1
1993s1
1990s2
1991s2
1992s2
1993s2
1990s3
1991s3
1992s3
1993s3
1990s4
1991s4
1992s4
1993s4
Proportion
Age (yr)

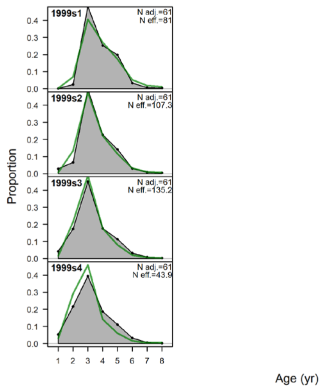

1994s1
1995s1
1996s1
1997s1
1994s2
1995s2
1996s2
1997s2
1994s3
1995s3
1996s3
1997s3
1994s4
1995s4
1996s4
1997s4
Proportion
Age (yr)


1998s1
1999s1
1998s2
1999s2
1998s3
1999s3
1998s4
1999s4
Proportion
Age (yr)

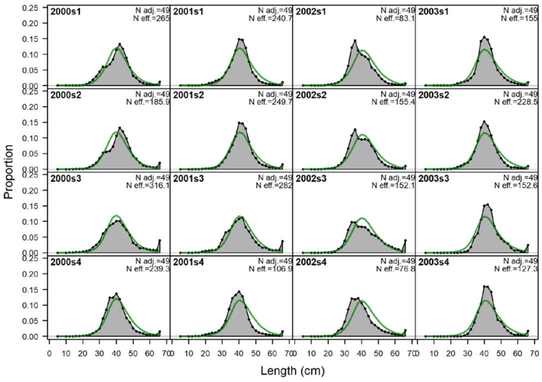

1987s1
1988s1
1989s1
1990s1
1987s2
1988s2
1989s2
1990s2
1987s3
1988s3
1989s3
1990s3
1987s4
1988s4
1989s4
1990s4
Proportion
Age (yr)

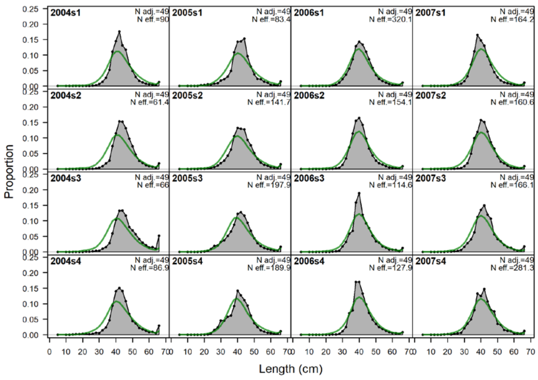

1991s1
1992s1
1993s1
1994s1
1991s2
1992s2
1993s2
1994s2
1991s3
1992s3
1993s3
1994s3
1991s4
1992s4
1993s4
1994s4
Proportion
Age (yr)

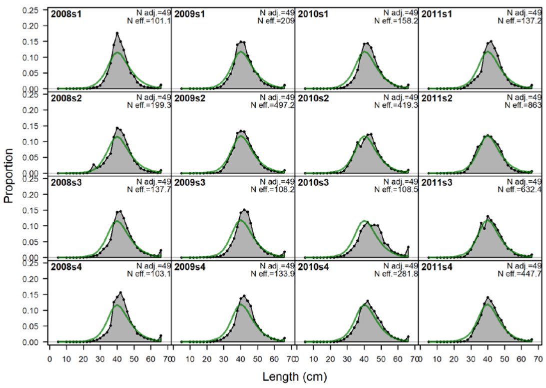

1995s1
1996s1
1997s1
1998s1
1995s2
1996s2
1997s2
1998s2
1995s3
1996s3
1997s3
1998s3
1995s4
1996s4
1997s4
1998s4
Proportion
Age (yr)

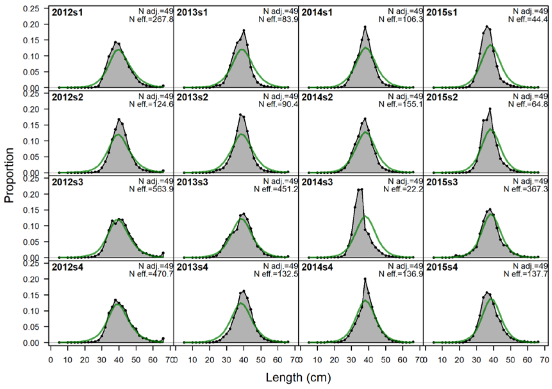

1999s1
1999s2
1999s3
1999s4
Proportion
Age (yr)

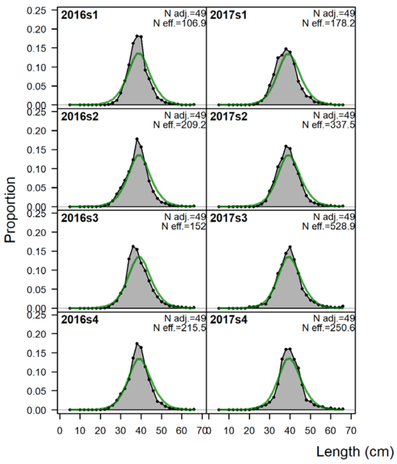
Fits to length compositions: Active, Passive, BITS Q1, BITS Q4

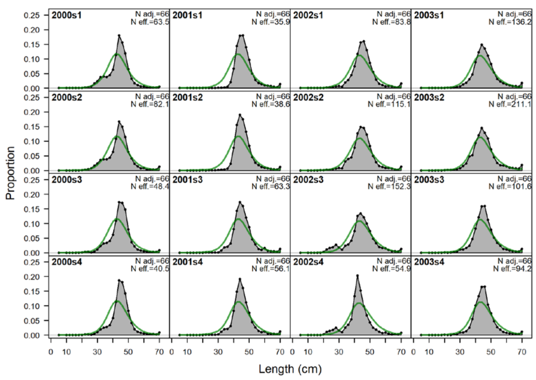


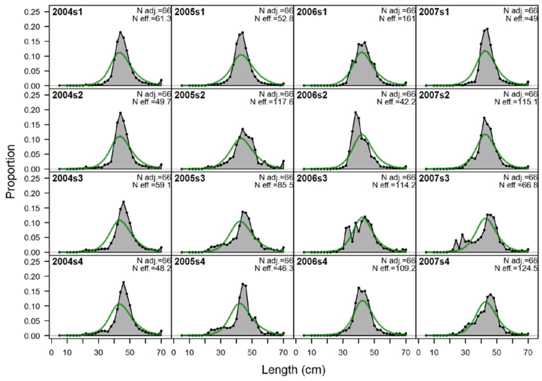


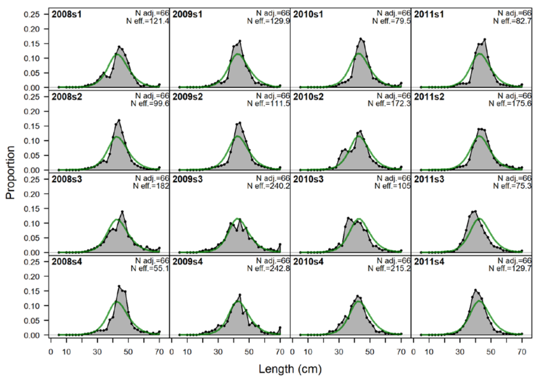


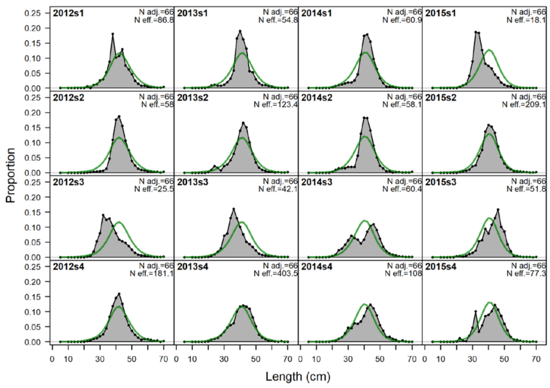


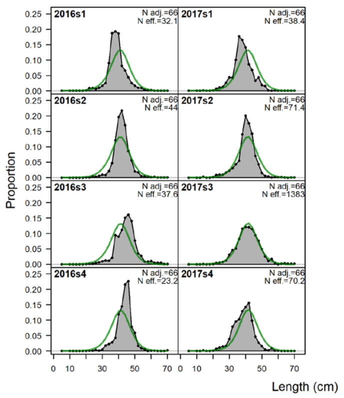


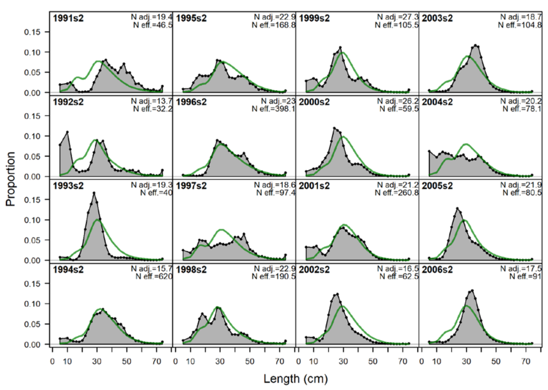

Proportion
Length (cm)

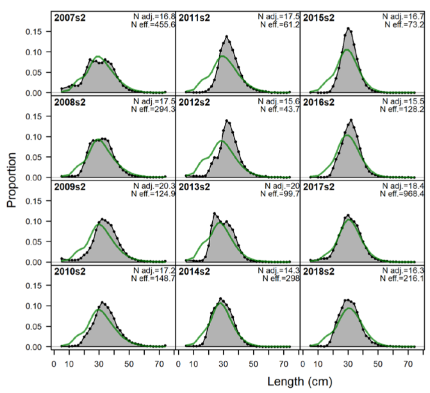

Proportion
Length (cm)

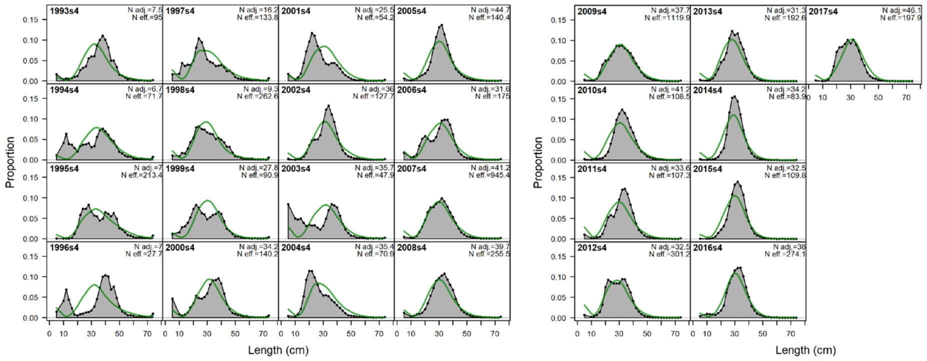

Proportion
Length (cm)


Proportion
Length (cm)

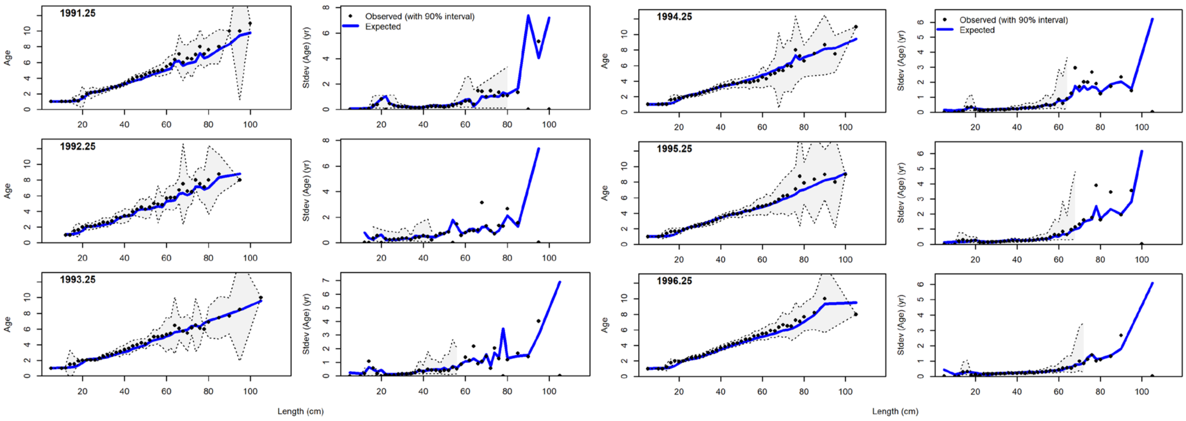

Proportion
Length (cm)

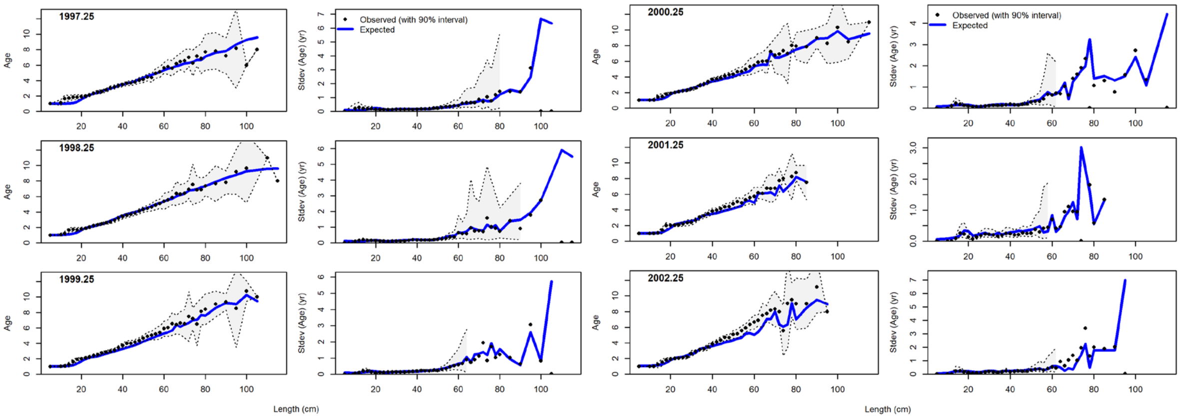

Proportion
Length (cm)

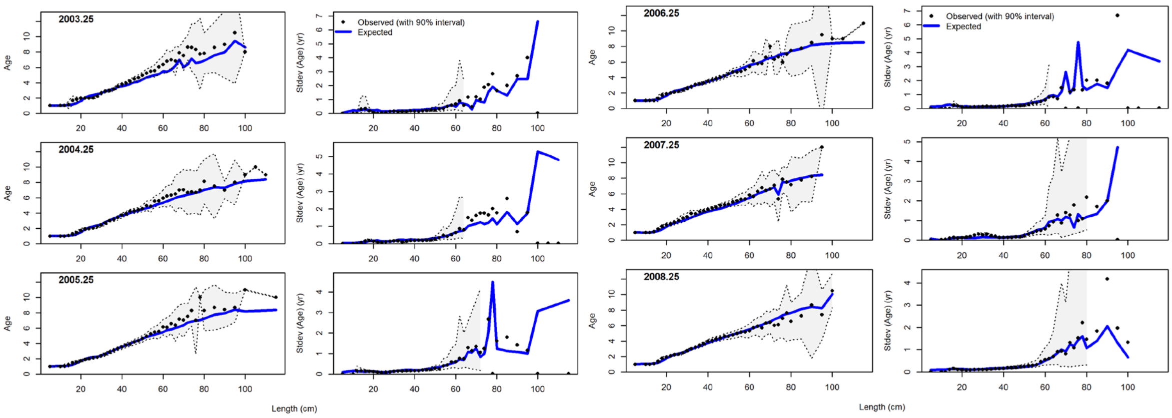

Proportion
Length (cm)

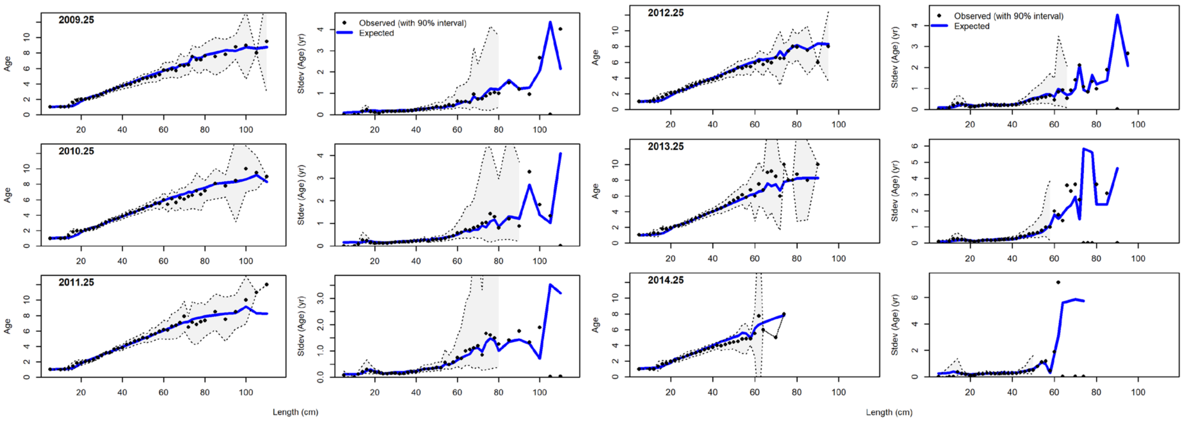

1993s4
1997s4
2001s4
2005s4
1994s4
1998s4
2002s4
2006s4
1995s4
1999s4
2003s4
2007s4
1996s4
2000s4
2004s4
2008s4
Proportion
Length (cm)
2009s4
2013s4
2017s4
2010s4
2014s4
2011s4
2015s4
2012s4
2016s4
Proportion
Length (cm)

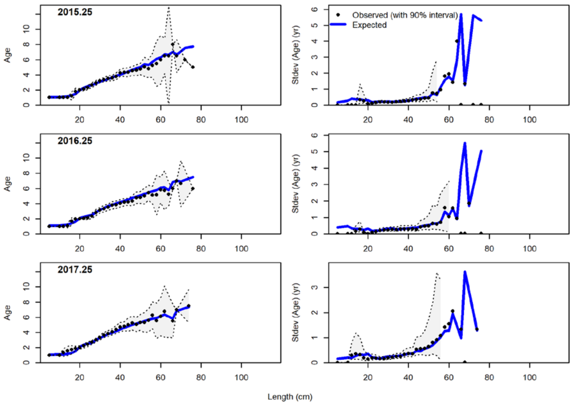
Fits to ALK: BITS Q1, BITS Q4

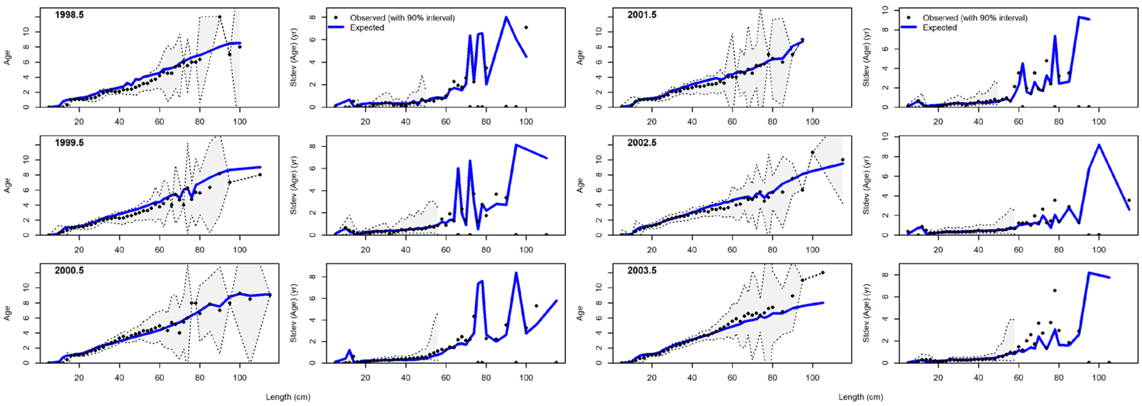


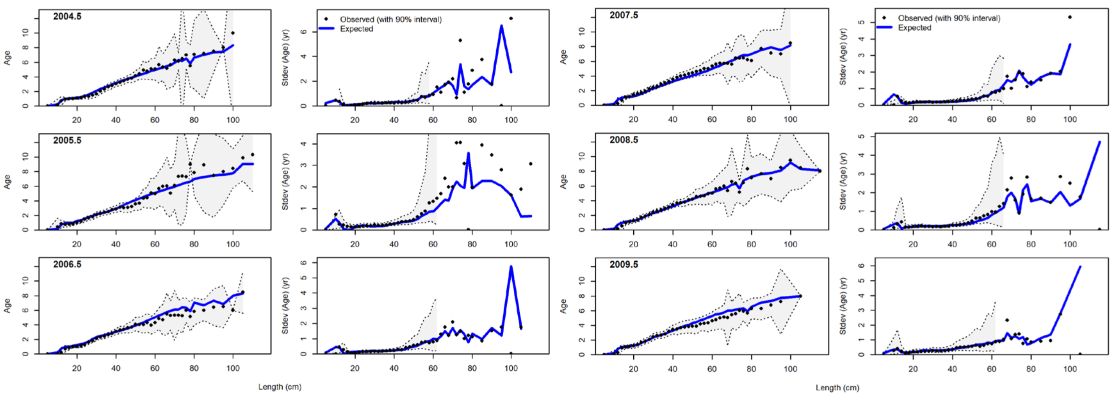


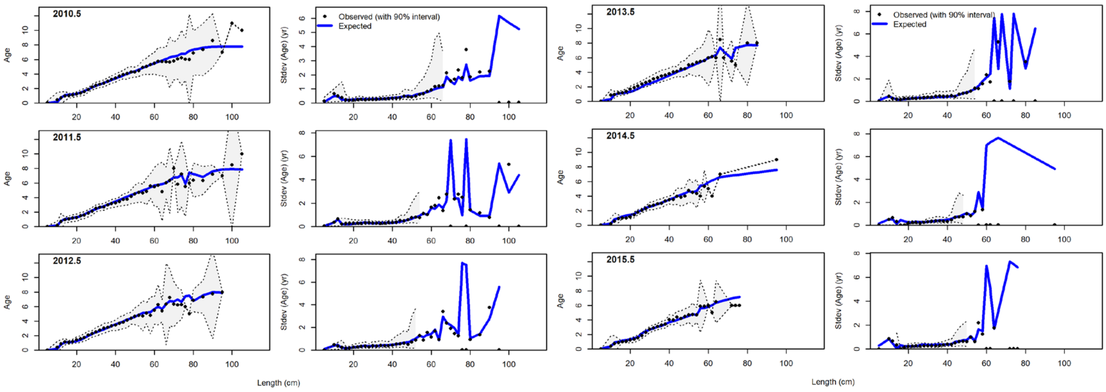


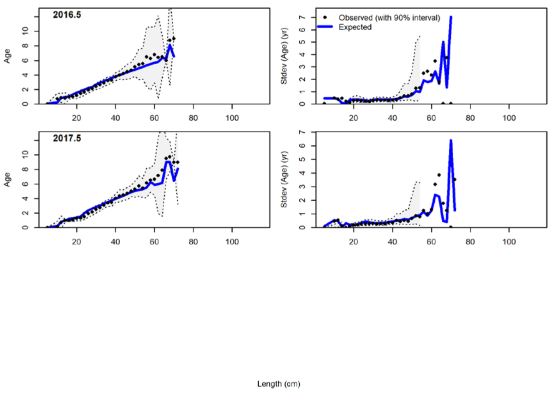

2009.25
2010.25
2011.25
2012.25
2013.25
2014.25
Observed (with 90% interval)
Expected
Age
Stdev (Age) (yr)
Length (cm)


2015.25
2016.25
2017.25
Observed (with 90% interval)
Expected
Age
Stdev (Age) (yr)
Length (cm)


1998.5
1999.5
2000.5
2001.5
2002.5
2003.5
Observed (with 90% interval)
Expected
Age
Stdev (Age) (yr)
Length (cm)


2004.5
2005.5
2006.5
2007.5
2008.5
2009.5
Observed (with 90% interval)
Expected
Age
Stdev (Age) (yr)
Length (cm)


2010.5
2011.5
2012.5
2013.5
2014.5
2015.5
Age
Stdev (Age) (yr)
Observed (with 90% interval)
Expected
Length (cm)


2016.5
2017.5
Age
Stdev (Age) (yr)
Observed (with 90% interval)
Expected
Length (cm)

WD8 - Workshop on use of recreational fisheries catch data in stock assessment of WB cod

17-19 October, DTU Aqua, 2800 Lyngby, Room 043, building 201, Kemitorvet, Denmark

Start: 1 pm on the 17th Oct; finish: 17.30 pm on the 18th Oct.

Main aims of the WK:

i) Follow up on the status and quality of recreational catch data in SD 22-24 in DE, DK, and SWE
ii) Move on the process of including DK and SWE recreational catches in stock assessment of WB cod
iii) Make a future workplan regarding recreational catch data for WB cod

Topic	Specific contribution/discussion item	Responsible
Recreational data in SA	Presentation of an overview how recreational catch data are used in stock assessment, incl. which are the important parameters for which information is needed (tons, size comp etc, length of the time series etc).	Rie (DK)
Overview of recreational catch sampling, by country	Present an overview addressing : i) methods how are recreational catch data presently collected and raised (both for tons and size/age comp) ii) overview of sampling intensity (how many trips, sites visited, how many individuals measured, for how many years and which type of data are available) iii)Which fleet segments involved in recreational fisheries are covered by sampling iv)Point out the weaknesses in present sampling, what could be done better v)Which factors are known to influence the inter-annual variations in the magnitude of recreational catch, i.e. are the estimates considered realistic?	Harry (DE) Nuno (SWE) Hans (DK)
Quality of the estimates for the magnitude of recreational catch, by country	i)Indications of uncertainties in the amount of recreational catch (variability between trips, sample sites etc). ii)Comparison of the estimated tons obtained from off-site and on-site surveys, where both are available.	Harry (DE) Nuno/ Andreas (SWE) Hans (DK)

	iii)Any indications, how do the regulations implemented affect the data on recreational catch in each country. Discuss and Conclude: **for which years will sufficiently reliable estimates for the magnitude of catch (tons) be available by January 2019, from SWE and DK**	 All
Assumptions on the historical magnitude of recreational catch	Which assumptions are made on the magnitude of DE recreational catch in years where no sampling is available? What is the evidence that these assumptions are reasonable? Discuss and conclude: **how to make reasonable assumptions on the magnitude of historical catch back in time, before sampling started in SWE and DK. What should the assumptions be based on?**	Harry All
Length distribution of commercial catch, by country/sub-region	Each country presents available information on length distribution of cod in recreational catch separating between SDs 22, 23 and 24. How does the length comp differ by quarters, by fleet segments etc? **Compare and conclude: where are length distributions similar, so that the data can be borrowed (btw which countries in which SDs, in which quarters)** Which assumptions are made for length distribution of DE recreational catch, in years where sampling is not available? Discuss and conclude: **for which years/sub-areas/ catch at length estimates for DK and SWE can be obtained by January 2019, based on recreational catch sampling, including possibilities to borrow data and make reasonable assumptions (what should these be based on?)**	Harry (DE) XX (SWE) Hans (DK) All Harry All
Age distribution of recreational catch	Where is the age information presently coming from, used to derive DE recreational catch at age Age information collected in recreational sampling in SWE and DK, how does this compare with commercial or survey (ALK)	Harry Hans (DK) XX(SWE)

	Conclude: **how to convert SWE and DK recreational catch at length to catch at age**.	All
Sum op of the status of data	**Conclude :** **i) What is available and where are the holes in DK and SWE recreational data that are presently not possible to fill by borrowing or reasonable assumptions** ii) **Is it possible to fill some of these holes in future, how and when?**	All
Future sampling of recreational catch	Coordination of future sampling, how do we sample most efficiently, so we can best borrow information across countries	All
Data storage	Possibilities to store recreational data in Intercatch, realistic time frame when will this be possible? How do we best store and exchange the data until IC can be used You are right that I have talked with Kieran Hyder and Estanis Mugerza about getting the recreational data in to the RDBES, and the plan is that the recreational data will get 2 or 3 specific data types in the RDBES for the data, which for now are different that the commercial data types. But we will first have time to develop the recreational data types around spring - summer 2019. So at this point there is not much to say other that what I just wrote. I have been give the data they want in the RDBES, so at this point we need to develop the RDBES for the commercial stocks first.	Rie invites a person from ICES Data Centre All
Other issues	Origin of cod in recreational catch from SD 24. Revisit the assumption of it all being western cod, can we get solid evidence for that in future, e.g. from genetics?	Harry
Workplan	**Prepare and agree on a workplan until January 2019 and longer term**	All

Participants

Participant	Participation	Nation	email
Joakim Hjelm Sweden		Sweden	joakim.hjelm@slu.se
Esha Mohamed		Sweden	esha.mohamed@slu.se
Nuno Prista	Skype	Sweden	nuno.prista@slu.se
Andreas Sundelöf	Wedensday	Sweden	andreas.sundelof@slu.se
Harry Strehlow		Germany	harry.strehlow@ thuenen.de
Uwe Krumme	Wedensday	Germany	uwe.krumme@thuenen.de
Simon Weltersbach		Germany	simon.weltersbach@thuenen.de
Hans Jakob Olesen		Denmark	hjo@aqua.dtu.dk
Marie Storr-Paulsen		Denmark	msp@aqua.tu.dk

Short minutes

It is important to keep in mind that the Benchmark is in 2019 (January / February) and therefore the Data compilation workshop will be in the fall 2018.

We have decided to have 2 intermediate meetings before the compilation workshop to follow up on decisions taken to this meeting.

1. Meeting will be a 1 day meeting before the WGBFAS. Presently we do not know if that will be hosted in ICES or in France.
2. Meeting will be in combination with the WGRFS in June where Harry, Simon, Andreas and Hans will participate

Main message on how data is presently used in the stock assessment.

As we do not use the recreational data as a tuning fleet, we cannot use a truncated time series and need the information on all age groups (presently 1994-present year) and age 0-7+. If data were to be used as a tuning series we could decide only to use a shorter time series and not all age groups. When used in stock assessment we will need CANUM (numbers of cod by age and year) and WECA (weight of cod by age and year). We would like to have an overview of sampling intensity (PSU) and numbers of fish.

Decision by the group
The data will be included in the future data call – under the assumption that the data and assessment has been approved at the benchmark in 2019. As data presently cannot be included in the RDB (maybe in 2019), data shall be sent as an excel file to Marie (weight and numbers by year and nation – WECA, CANUM and total weight) and Marie will compile the data sources from all 3 countries.

Sampling program

Denmark

Presently Denmark only samples biological information from charter boats in SD 23 and not SD 22 (2-3 charter boats).

A suggestion was to make a pilot on the charter boats in SD 22 in 2018 and compare the results (length distribution) with the tur boats information from Germany. To be followed up by Hans and Simon and presented at the WGRFS in June 2018

The calculation of the total estimate is based on the mean catch per angler per trip which is multiplied by the total number of trips for each vessel per quarter. The number of trips for each vessel is not estimated but is census data. The number of visits to the particular vessel is still fixed and the vessel will be visited on the succeeding trip with an equivalent day-type (weekdays or weekend).

It should be kept in mind that 0 observations are not included if only charter-boat trips conducted are sampled. Another design could be to have random days (week day and week-ends as now) and then a charter-boat is called and if a trip is not conducted, due to bad weather or other this is a 0 trip and should be registered. In this system refusal could also be included.

The Danish preliminary onsite results from SD 23 indicate that the CPUE were larger (nearly double) of the value obtained by the DST (Statistic Denmark) estimate for the same year. Since 2009, Denmark has information from DST on recreational catches (numbers or weight) by SD and if the fishery was by passive gear or angling. There is no detailed information on effort related to areas or species in this telephone/internet survey. The survey is conducted bi-annually and participants are asked to give data on their catches for each quarter of the year. Around 4000 license holders are participating annually (response rate 40 – 50 %).

A suggestion was to decrease the numbers of license holders in the survey with 50% and instead make the survey interview 4 times a year. The bag-limit introduced in 2017 has anyway changed the time series for cod catches and hence it could be the time to change.

Effort is presently only obtained from the omnibus survey from 2010. It would be very good to confirm the effort (number of days per year or week and platform a fisherman with a license is fishing).

To compare the charter-boat length distribution in SD 22 (pilot – not preformed yet) with the German length distribution in SD 22.

Sweden

Sweden has been collecting catch and effort data on tourboats for some years now. In 2017 such information has being extended with in situ onboard observations of catches (including releases), lengths and biological information (e.g., otoliths for age determination).

In what concerns the private boat component Sweden does not have a register of either vessels or fishers. Such absence of register significantly limits the usefulness of off-site survey methods like mail or phone surveys that have to rely on less efficient sampling frames[1]. As a consequence, a pilot field-survey aiming to collect in situ data for effort and catch estimation is been trialed in 2017.

The pilot field-survey is carried out quarterly in subdivisions 23 and 24 (considered strata). In each quarter, a set of sampling dates per subdivision are distributed systematically (with random start) into 6 sampling waves. The design is stratified multi-stage with date as primary sampling unit (PSU), municipality as SSU, workshift as TSU and harbour as QSU. The sampling frame for PSU includes all calendar dates and both weekends and weekdays are sampled. The sampling frame for SSU includes all kommunes in each subdivision plus 1 fringe kommune with slightly lower probability (Ängelholm, Simrishamn). Sampling frame for TSU includes three shifts: 06-14, 14-22, 22-06. Slightly less probability (p=0.2) was assigned to the 22-06 shift where effort and catches was expected to be lower but all hours have non-zero probability. All recreational fishing ports and ramps have some probability of being selected. In practical terms, in each day a couple of observers go to a randomly drawn municipality in a randomly drawn workshift. Then they carry out a bus route access-point survey among a subsample of randomly drawn access points spending a fraction of time in each port where they interview incoming vessels. In most municipalities all ports are sampled; in Höganäs, Helsingborg, Malmö – where access points are many – a subsample of access points is selected using an algorithm that, while maintaining a considerable degree of randomization, is able to maximize observation time by ensuring that sampled sites do not distance more than 20-30 min from each other. The previous methodology secures that interviewing time in each municipality is between 5H15 (min) and 7H40 (max) in each 8hr shift, effectively maximizing both the numbers of access-points visited and the interviewing time.

The design used to quantify effort and catches of private boats assures a very large number of PSUs and is able to deliver (so far judged) unbiased estimates of both effort and catches. Also, in SD 24 (where sampling effort has been lower) very low numbers of actual interviews have been conducted so far: such situation may be a consequence of significantly lower effort and catches in that stratum but also a consequence of too low a sample size.

The survey design allows for calculation of the effect of time of the day on CPUE by stratifying sampling in 3 time slots. It also circumvents the lack of a priori information on the most important access points by surveying with known-probability all existing access-points. Some doubts subsist with regards to the precision of the estimates and the capability of current effort levels to detect the signals generated in the fishery. The effect on total estimate of trips and catches when increasing the effort with 100% in SD 23 (from 12 to 24 days in Q3) was analyzed and did not indicate a substantial difference. Presently there is biological sampling from private boats but, similarly to other countries, the number of fishermen interviewed with cod catches is insufficient to draw a separate length or age frequency for the private boat component. Age information from other components (e.g., tourboat, commercial) may therefore have to be used when providing information for stock assessment.

A suggestion was to increase the sampling effort in SD 24. Or by stratification make sure data is obtained. It was pointed out that the low deviation between the two sampling intensities in Q3 may be completely

[1] Note: Despite inefficiency Sweden does have a country-wide mail survey on-going for some years. That survey is the basis for estimates of several stocks in several bodies of water (including inland waters) but does not, unfortunately, have the spatial and temporal resolution required by WB cod assessment.

random and need to be verified over a longer period of time. Equally important to a robust statistical design is knowledge of the fishery. By ensuring that all access-points and time periods are being sampled the present design is slowly collecting such information. By the end of 2017 some adjustments may be made to increase the efficiency of the design and provide for a larger chance of interviewing cod fishers.

To compare the CPUE from before the implementation of the bag-limit and after the implementation. Contact harbor masters, fishing guides, angling clubs to obtain present and historical catch data.

To compare the length distribution between Swedish and Danish charter-boats in SD 23, if length distribution is not considered significant different a combined sampling program with all charter-boats in one frame could be considered and depending on the harbor site appointed to either Sweden or Denmark.

Sweden has introduced a wave sampling system were all harbors with in SD 23 and 24 are included and some harbors close to the border on both sites are included as well. In this bus-route access-point survey a technician is sent on a trip to harbors within a municipality randomly selected. Sampled sites are within a 20 min distance of each other. Each site is visited for 50 min. This design assures a very large number of PSUs but has made it difficult to obtain biological information. From SD 24 very little information has been recorded. The survey design allows for calculation of the effect of time of the day on CPUE by stratifying sampling in 3 time slots. The effect on total estimate of trips and catches when increasing the effort with 100% in SD 23 (from 12 to 24 days in Q3) was also presented but did not show a significant difference. Presently there is no biological sampling from private boats.

A suggestion was to increase the sampling effort in SD 24. Or by stratification make sure data is obtained.

To compare the CPUE from before the implementation of the bag-limit and after the implementation

To compare the length distribution between Swedish and Danish charter-boats in SD 23, if length distribution is not considered significant different a combined sampling program with all charter-boats in one frame could be considered and depending on the harbor site appointed to either Sweden or Denmark.

Germany

The German time series goes from 2002 to present time but data is considered reliable from 2005. The effort information is obtained from the diary survey dating back in time. Presently Germany has a high number of on-site visits on the different platforms (on-shore, charter-boats and private boats). Length information is obtained from tour-boats and used for private boats and on shore. In general, biological and catch data is stratified into boat-based and land-based fishing activities. No weight or age information is obtained from the program but is borrowed from the commercial fishery by SD and half year. In later years the importance of the private boats sector has increased while at the same time been decreasing for the charter-boats. This trend seems to have been enhanced by the bag-limit making the hinterland ☺ more reluctant to travel to the sea. The charter-boats fishing in SD 23 is presently not covered by the sampling program, and the catches are landed in Denmark.

A suggestion was to compare the German effort from the on-site survey with the mail diary survey.

It seems like a fair approach to borrow length data from charter-boats to private boats but it would be good to test the length distribution from land-based angling with length data from the charter boats to see if they are indeed similar.

To compare if the CPUE is different on different times of the day (morning / afternoon)

Denmark and Germany could compare the fixed effort used from the Danish Omnibus and the German mail-diary survey with numbers of days fished when obtaining an annual license or weekly license .

To archive effort information from charter boats fishing in SD 23, then a Danish or Swedish length distribution and CPUE could be applied.

Germany could compare the index for the BITS and SSB (from the assessment) with the tons in the recreational fishery

General comments

It could potentially be a problem with the German and Danish fixed effort used (days fishing when holding an annual / weekly license) although they are in both countries weighted by the annual license.

The land based post release mortality should be investigated. Presently we are using a 100% mortality and a 11.2 % mortality on sea based recreational catches (Ferter et al., 2015, Weltersbach and Strehlow, 2013, Capizzano et al., 2016).

Sweden was considering having an overflight survey in SD 23. Demark has an on-site roving creel boat survey, could be combined if the Danish boat is also charting Swedish anglers in SD 23. Uncertainties in peak fishing activity and difficulties in randomizing flight times (e.g., under different weather conditions) should be considered when exploring such alternatives.

To be able to compare estimates between countries, we need to be aware of the effective sample size. The effective sample size consists of the design effect and the gross sample size. The design effect gives information of how much the precision of an estimate is increased or decreased by using a complex design instead of a simple random sampling design (baseline). The effective sample size should be fixed and agreed upon countries where each country should aim to achieve.

Other issues that need to be put into consideration is using a probability sampling design, check if there is any under-coverage (both temporal and spatial) and the level of non-response.

Bag-limit

Germany do not seem to have any problems with the limitation as the CPUE is very low presently and therefore it is very seldom an angler is above the limitation. However, fishing effort in particular on charter boats has decreased noticeably.

Sweden has experienced a changed behavior where anglers onboard a charter-boat is sharing the fish if catching more than the bag-limit. This is probably against the intention of the regulation (as it potentially does not limit the fishery). However, although compensation occurs the overall harvest may still be reduced due to the bag limit (if in average all the anglers on board are catching more than 5 fish a day) . Sweden has experienced many skippers on charter-boats who are enforcing the regulation and has increased the minimum landings size to 40 or 45 cm on a voluntary basis (this will only be an improvement if the post release survival is high). Sweden has experienced that the effort i.e. numbers of participants on charter-boats has decreased but not the total numbers of trips.

Denmark has also experienced a changed behavior where anglers onboard a charter-boat is sharing the cod if necessary to get below the bag-limit. There is no information on enforcement of the bag-limitation by skippers and it has been shown that charter-boat anglers sometimes do not comply with the bag-limitation.

Decision by the group
As a minimum requirement the group agrees that we need a length distribution by SD and by fleet segment (passive gear fishing / angling from sea) for the use of stock assessment. Age and weight would be very good to have but could be borrowed until the information is obtained from same sampling source.

Extrapolation back in time

Data needs to be extrapolated back to 1994. We agree on a prioritization:

1. Charter-boat
2. Private boat
3. Passive gear fishing
4. Land-based

We would like to have both Danish and Swedish recreational data for charter-boats and private boats but passive gear and land-based fishing fishing will have a lower priority.

Land-based fishing activities may be substantial and although in the case of Germany they are only responsible for 15% of the total catch in weight, the contribution in total numbers could be higher due to the smaller length classes caught.

To reconstruct the time series we need to contact harbor masters, fishing guides, angling clubs to obtain historical catch data

Decision by the group
We would like to include catch-in –number and catch-in-weight by age from the timer period 1994-present from the Danish and Swedish Charter-boats and private boats in SD 22-24 for the assessment – data to be ready fall 2018.

For charter-boats, the length distributions in SD 22 (Danish pilot study and the German data) needs to be compared. If not significantly different, the German length distribution will be applied to the Danish charter-boats and private boats recreational catch data from SD 22 (2009 present time). Before 2009 we do not have Danish data on total catches (tones) and another approach needs to be used.

Simon has 9 years of data with length distribution from SD 22.

Further back in time the length distribution could be obtained from the BITS survey together with the 9 years of German charter-boat survey. If we can model the length distribution back in time using the BITS length. If we are not able to use the survey data (as they could miss out some of the larger fish), we will need to use the commercial data compared to the recreational German data (modelled). We need to find a person that would like to model the length distribution back in time.

Decision by the group
We would like all three countries to use the same approach back in time for length distribution when no real data is available. This indicate that we would change the German time series before 2010. Presently Germany is using 2010 length distribution back in time.

Other issues

Decision by the group
For recreational fisheries we agree that we will not use the split (maybe at next benchmark). And will therefore not include genetics or otolith shape for split

For the short term forecast what are we doing with the intermediate year (partial F compared to Fixed)

References

Capizzano, C.W., Mandelman, J.W., Hoffman, W.S., *et al.* (2016) Estimating and mitigating the discard mortality of Atlantic cod (*Gadus morhua*) in the Gulf of Maine recreational rod-and-reel fishery. *ICES Journal of Marine Science: Journal du Conseil* **73, 9**, 2342-2355.

Ferter, K., Weltersbach, M.S., Humborstad, O.B., *et al.* (2015) Dive to survive: effects of capture depth on barotrauma and post-release survival of Atlantic cod (*Gadus morhua*) in recreational fisheries. *ICES Journal of Marine Science* **72, 8**, 2467-2481.

Weltersbach, M.S., Strehlow, H.V. (2013) Dead or alive—estimating post-release mortality of Atlantic cod in the recreational fishery. *ICES Journal of Marine Science: Journal du Conseil* **70**, 864–872.

WD9 - Workshop on separating Eastern and Western Baltic cod for stock assessment

3-4 July 2018, Lyngby, Denmark

Participants	Mail	Country
Uwe Krumme	uwe.krumme@thuenen.de	Germany
Franziska Schade	franziska.schade@thuenen.de	Germany
Henrik Mosegaard	hm@aqua.dtu.dk	Denmark
Christoffer Moesgaard Albertsen	cmoe@aqua.dtu.dk	Denmark
Julie Davies	joco@aqua.dtu.dk	Denmark
Karin Hüssy (Tuesday)	kh@aqua.dtu.dk	Denmark
Margit Eero	mee@aqua.dtu.dk	Denmark
Marie Storr-Paulsen	msp@aqua.dtu.dk	Denmark
Jakob Hemmer Hansen (by skype)	jhh@aqua.dtu.dk	Denmark

Main aims of the Workshop:

i) Standardize and agree on methodologies used for shape analysis
ii) Include samples from both DE and DK commercial data for splitting commercial catches in SD 24
iii) Split survey indices in SD 24
iv) Coordinate future sampling and work distribution between DE and DK

Meeting center Ly202-R1020 Kemitorvet ; Thuesday at 9.15 -18.00- Wedensday 8.30 – 14.00

I Methods

OTOLITH SHAPE	
Shape analyses method	Henrik/ Christoffer: present the i) shape method used so far for splitting DK data, and how is baseline used, ii) shape analyses results with test set including known origin Franziska: present the i) present the shape method used for splitting by DE (the R package) (e.g., smoother, bias correction, covariance etc), and how is baseline used ii) shape analyses results with test set including known origin

Which factors to account for	Karin: present analyses on significance of i) quarter, ii) length-group, and ii) sub-area within SD 24, for split, based on DK commercial data. DE(Franziska?): Present analyses on significance of i) quarter, ii) length-group, and ii) sub-area within SD 24, for split, based on DE commercial data and/or survey data.
Conclusions on methods	All: Discuss and conclude on i) The shape analyses method to be used in future. ii) Necessity and possibilities for updating the genetic baseline. How to apply the baseline for splitting survey, vs. commercial data, can the same baseline be used for both? (Internal vs external baseline) iii) which variables should be accounted for in split (length, position, quarter)

II application of split

SURVEY:	
Coverage and sample selection for split (otolith shape) analyses	All: Germany has earlier presented the output from the split on the survey – do we need to modify based on the conclusion on agreed method (I), Discuss and agree on scheme for selection of otoliths for shape analyses from surveys, incl. whether or not both Q1 and Q4 should be analysed on a routine basis. To assign small cod <20 cm in the survey. Denmark has genetic for small cod (see manuscript).
Applying output from split analyses	All: Agree on the format for data delivery to survey index calculation
COMMERCIAL:	
Coverage and sample selection for split (otolith shape) analyses	All: Discuss and agree on the scheme for selection of otoliths for shape analyses from commercial catch, by DE and DK. Do we keep the present split by area (I and II)

Minutes from the meeting

Baseline

Present baseline used for the Danish commercial split is from 2011 and 2012 and is composed of 971 genetic samples and additional 1969 spawning samples in the length range from 20 – 106 cm. Present baseline used for the German split is from 2015 (?) and is composed of 507 samples in the length range from 26 – 79 cm.

It would strengthen the baseline if more genetic data to be included in future baseline 2003, 2004 and 2014 (used for the cross validation for this meeting). Further, it seems like Jakob has used more than 600 (extra) fish (juvenile fish) in his paper were we presently do not have the shape included in our baseline.

- Jakob is sending a list to Karin, with the individual fish ID, survey and year (by 10. of August latest). Karin is matching all the fish with otolith pictures and where there is a picture missing Karin is sending a mail (latest the 15 august) to Julie/ Uwe asking if the otolith can be found, photographed and included in the base line.
- New baseline should be finalized by the 1. September. Christoffer will merge the new baseline. But at the time he will need to look into if we can keep the 4 length group or if this has to be changed to a continual length group / more length groups or stick to the present method.
- The pictures from the baseline need to uploaded in 3 new folders:
  1) juvenile fish – this has the highest priority as it will be needed for the German suvey split - proxy 600 fish (Casper Berg will need this data to finalize the new survey index before the benchmark)
  2) A updated Danish baseline including 2003, 2004 and 2014 – if new pictures have been found. Further, Denmark will clean the baseline for “bad” pictures and update the baseline with new samples 2003, 2004 and 2014. (possible more pictures if we will have the time)
  3) The German baseline (is already available).
- Updates to present baseline for benchmark:
  DK: including 2003, 2004 and 2014 and the juveniles fish data
  DE: include juvenile fish data

Genetics

- Franziska to contact Peggy (latest mid-July) and contact Jakob to investigate and document the statistical power of the German genetic method (by similar analyses as done by Jakob for Danish data). Report back the conclusion of this investigation to the WK attendees 1 September.

- If the otolith shape and genetics are equal then the fish is assigned to the genetics signature no matter where it was caught and when it spawned. (Germany) – this will be applied in Denmark as well.

Length

- Assignment by length group – presently Denmark is using 4 length group this could be investigated if the groups are appropriate – maybe it should be a continues length or other groups
- German samples should be standardized according to length 1. Residuals 2. Stock length (slope can be different).

Quality of pictures

- German pictures need to have in the name if it is a left or right. (for the new small cod)
- Crystalline otoliths need a mark as well as glued and not glued, and dirty
- Convex side up when pictures are taken
- Check if a light from below can improve quality
- Denmark will go through all the pictures manually and make sure the quality is better(Julie)

Cross validation

At the meeting a cross validation was performed.

A test set with 489 samples with genetic information was used.

	German Baseline (Class. Success)	Danish Baseline (Class. Success)
German method (test set)	**36.3% W : 63.7% E (75%)**	**9.8%** W : **90.2%** E **(61%)**
Danish method (test set)	17.7% W : 82.3% E (61 %)	36.7% W: 63.3% E (75.2%)
True value (genetic)	**26.5% W :73.5% E**	**26.5% W :73.5% E**

Both methods performed equally well when the method and the baseline came from the same country but performed very bad when the baseline from the other country was used with own method. The reason behind this should be looked further into, but it was suggested at the workshop that it could relate to how the respective baselines were selected and applied in the two cases. The Danish baseline is much larger than the German although unbalanced with more small eastern cod and large western cod. The German baseline is 1/3 of the Danish however covering the same length classes for the 2 areas.

Roadmap

Short term:

Tasks	**Who and when**

Proportions of EB and WB cod in German commercial catches for the years available (separately for Area 1 and Area 2)	Franziska sends to Margit before 15 of august
German commercial landings by ICES square back in time	Rie checks whether we have, or otherwise contacts Sven (before August 15)- has been done
Splitting proportions from German historical survey (Icebeer) for all the years available (separately for Area 1 and Area 2), including only fish >30 cm	Franziska sends to Margit before 15 of august
Update the baseline with juvenile fish	DK and DE, finalized by Sept 1.
Update the regular baseline (dk)***	DK: should be finalized by Sept 1.
DE re-analyses the stock proportions in Solea survey with updated baseline (incl small cod)	Franziska sends the individual fish data to Casper (incl. fish length, lat, long, year, season, stock origin), both Q1 and Q4 for all the years available, by Oct 1.
WD (s) for ICES Data meeting documenting: DE: i) genetics and otolith shape methods; ii) the data coverage (how many fish, spatio-temporal coverage) for commercial catch, historical surveys; and BITS survey; iii) resulting proportions DK: updates the baseline and any updates to the shape analysis method? DE, DK: a common WD describing the cross-validation exercise of this WK	All by 1 Oct

***Denmark will try to change the base line size sorting group to fit the commercial size sorting groups. Presently the size sorting groups are <32, 32-37, 37-42 and >42.

Decisions concerning the methods to be applied in stock assessment at benchmark 2019:

BITS survey	Apply DE baseline with DE shape method (needs to be re-run with updated baseline incl. small fish)
Commercial catch	DE catch: split based on DE data/methods/analyses for the years available DK catch: split based on DK data/methods/analyses for the years available Other countries: split based on average of the available data Historical catches back to 1977: split based on DE methods/analyses, applied on DE survey

(longer timeframe)

- We would like to have a common baseline. Include the Swedish and Polish otoliths
- Investigate the recreational data for stock origin
- Further developments of the molecular tools to facilitate a shift to genetics based split.

Data to be send

The 2 analysis with German base line with German method with the test samples.

with Danish base line with German method with the test samples

and including the individual assigned otoliths. And the German confusion matrix (the 4 numbers)

Franziska to send to Henrik

WD 10 - Stock splitting of western and eastern Baltic cod in SD 24

Franziska Schade[1], Peggy Weist[2], Uwe Krumme[1]

[1] Thuenen Institute of Baltic Sea Fisheries, Germany, [2] Thuenen Institute of Fisheries Ecology, Germany

Background

Since the cod benchmark 2015 (WKBALTCOD), Baltic cod stocks in ICES subdivision (SD) 24 (Arkona Basin) are separated according to their stock assignment into western Baltic cod (WBC) and eastern Baltic cod (EBC). However, presently there are only annual mixing proportion estimates of WBC and EBC in SD 24 based on Danish commercial catches from active gear fisheries.
The German BITS survey in SD 24 is currently not included in the assessment, although the survey data have a wide spatial (all rectangles in SD 24) and temporal coverage (1977 to 2017).
Commercial catches from passive gear fisheries are also not taken into account, even if there is evidence that mixing proportions differ with water depth, i.e. large proportions of WBC in shallower waters are fished by passive fishing gears. To include the BITS survey and German commercial data in the stock splitting, an alternative stock assignment method was developed.

Genetics and otolith shape analysis were combined to establish a genetically validated otolith baseline with stock-specific shapes, providing the basis for an individual assignment of unknown fish otoliths to their stock of origin. This combined approach enables the quantification of present and past mixing proportions of Baltic cod stocks, originating from surveys and commercial catch sampling.

Material and methods

Sampling

To create a genetic and otolith shape baseline, cod were sampled along the southern Baltic Sea (SD 22-25) from commercial, survey and recreational catches (N_{total}=519, Table 1). Spawning individuals (maturity stages 5 and 6, Tomkiewicz et al. 2003) from the Belt Sea (SD 22) and the Bornholm Basin (SD 25) were used as reference samples for the western and eastern Baltic cod stock, respectively.

Table 1: Summary of cod sampling including capture area (ICES subdivision and rectangle) and period (year/month), sample size (N), total fish length (range and mean ± SD (standard deviation)), proportion of spawning individuals (stage 5 and 6), sex ratio (only females presented) and sample origin with fishing gear information.

Subdivision	Rectangle	Year/month(s)	N	Length range [cm]	Mean length ± SD [cm]	Spawning fish [%]	Female fish [%]	Sample origin (fishing gear)
22	37G0	2016/03	12	31-79	65.08 ± 14.51	100	25	Survey (bottom trawl)
	37G1	2016/02	26	40-77	59.04 ± 9.65	100	54	Commercial (gill net)
	37G1	2016/02+03	18	34-72	51.06 ± 11.89	100	39	Survey (bottom trawl)
	37G1	2016/07	1	44	44.00 ± 0	100	100	Commercial (bottom trawl)
	38G0	2016/03	1	65	65.00 ± 0	100	100	Survey (bottom trawl)
	38G1	2016/03	4	26-68	41.75 ± 18.95	100	50	Survey (bottom trawl)
23	40G2	2016/03	58	29-55	38.02 ± 4.83	22	28	Recreational (fishing rod)
24	37G3	2015/10	53	43-50	46.45 ± 2.00	0	57	Commercial (gill net)
	37G3	2016/05	55	42-50	47.07 ± 2.10	0	69	Commercial (gill net)
	37G4	2016/06	23	37-67	44.87 ± 6.84	9	57	Commercial (gill net)
	38G2	2016/04	57	49-58	53.47 ± 2.67	5	74	Commercial (gill net)
	38G3	2015/12	54	38-52	44.11 ± 4.06	0	48	Commercial (bottom trawl)
	38G3	2016/05	60	38-45	40.50 ± 1.80	17	73	Commercial (bottom trawl)
	38G4	2015/09	57	38-56	41.67 ± 2.61	0	75	Commercial (bottom trawl)
25	38G5	2016/02	11	30-39	34.45 ± 2.54	100	9	Commercial (gill net)
	38G5	2016/03	10	36-46	39.50 ± 3.89	100	70	Commercial (gill net)
	39G5	2016/05	10	34-49	40.30 ± 4.97	100	70	Commercial (gill net)
	39G6	2016/06	9	32-50	42.22 ± 5.47	100	22	Survey (bottom trawl)

Genetics

To differentiate between Baltic cod populations, diagnostic SNPs (Single Nucleotide Polymorphism) were selected based on whole genome-sequencing data from a total of 115 cod specimens originating from the North Sea (SD 4.b), Öresund (SD 23), Kiel Bight (SD 22), Arkona Basin (SD 24), and Bornholm Basin (SD 25). SNPs with high discriminatory power between populations were identified based on pairwise F_{ST}-values calculated for each locus, resulting in a panel of 23 SNP markers to unambiguously discriminate between western and eastern Baltic cod.

We then used this panel to evaluate the efficiency of our minimum SNP set by genotyping 519 individuals from SD 22-25 (Table 1). Low-performing markers were removed from the set, reducing the final set to 20 SNPs. Using the programme GeneClass2 (Piry et al. 1999) we assigned individuals to the most likely reference population (WBC or EBC) based on genotype likelihoods (following Rannala and Mountain 1997), which were used to calculate assignment scores. The distributions of likelihood ratios were well separated, corresponding to an assignment rate of 99.1% for which the assignment was unambiguous; only one spawning individual caught in SD 22 was assigned to the EBC stock (Figure 1 and 2). Individuals from SD 23 and the mixing zone SD 24 were assigned to the WBC or EBC stock using Principal Component Analysis (PCA) implemented in the EIGENSOFT v5 software (Patterson et al. 2006). Eigenvectors were inferred using only individuals from a subset of all samples (spawning individuals from SD 22 and SD 25), and then individuals were projected onto those eigenvectors to be assigned. The PCA revealed two distinct clusters (WBC and EBC) and individuals from SD 24 clustered with both, western and eastern Baltic cod, indicating mechanical mixing of both stocks in this area (Figure 3).

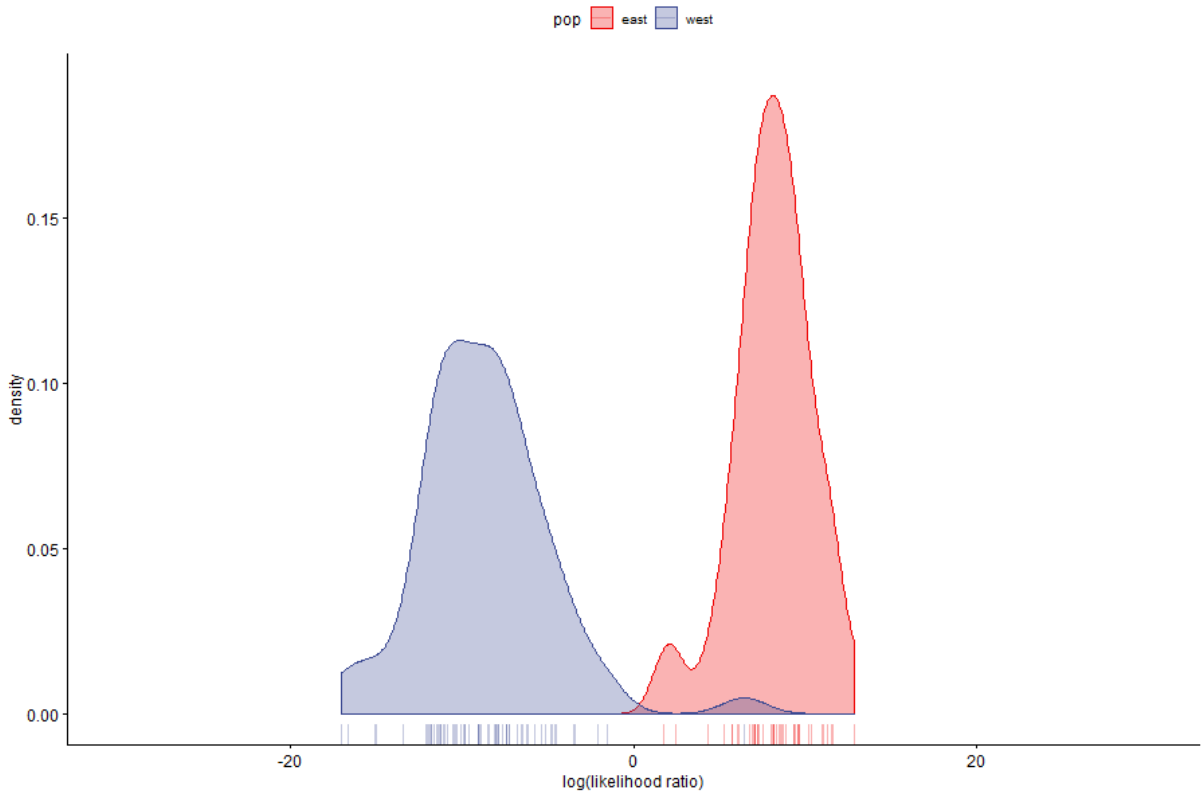


Figure 1: Distributions of Log(Likelihood ratios) for baseline samples from SD 22 (west) and SD 25 (east) collected in 2016 based on the minimum SNP panel (20 SNPs).

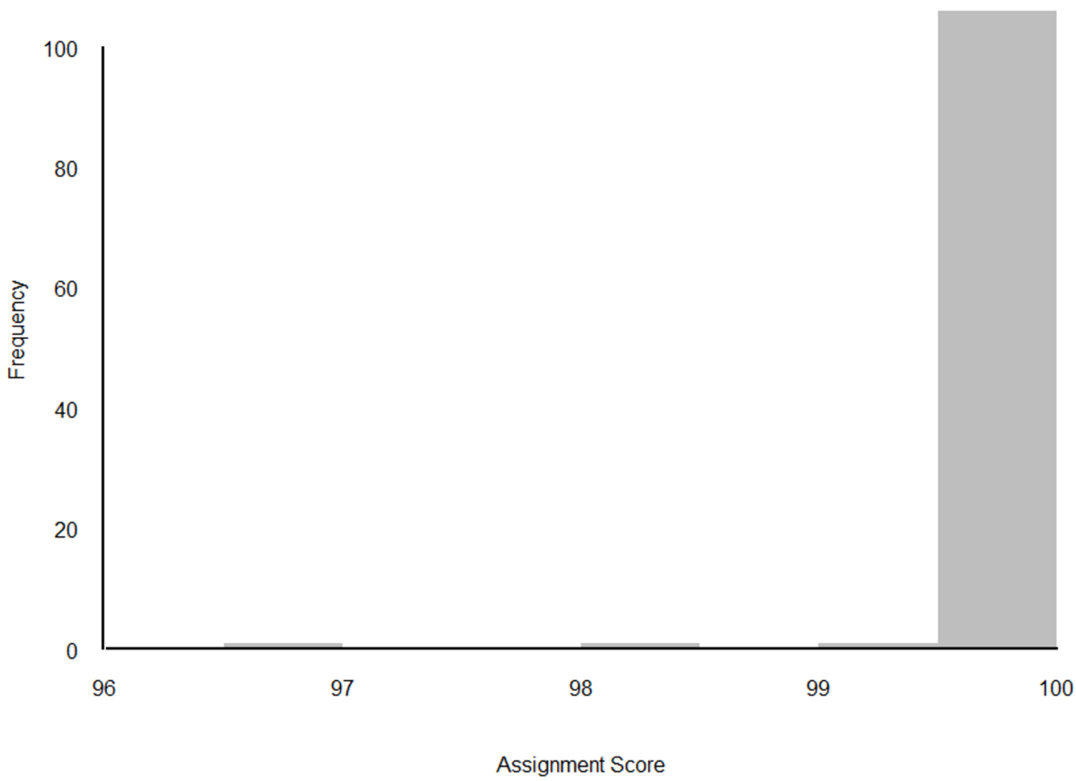


Figure 2: Distributions of assignment scores for baseline samples based on the minimum SNP panel (20 SNPs).

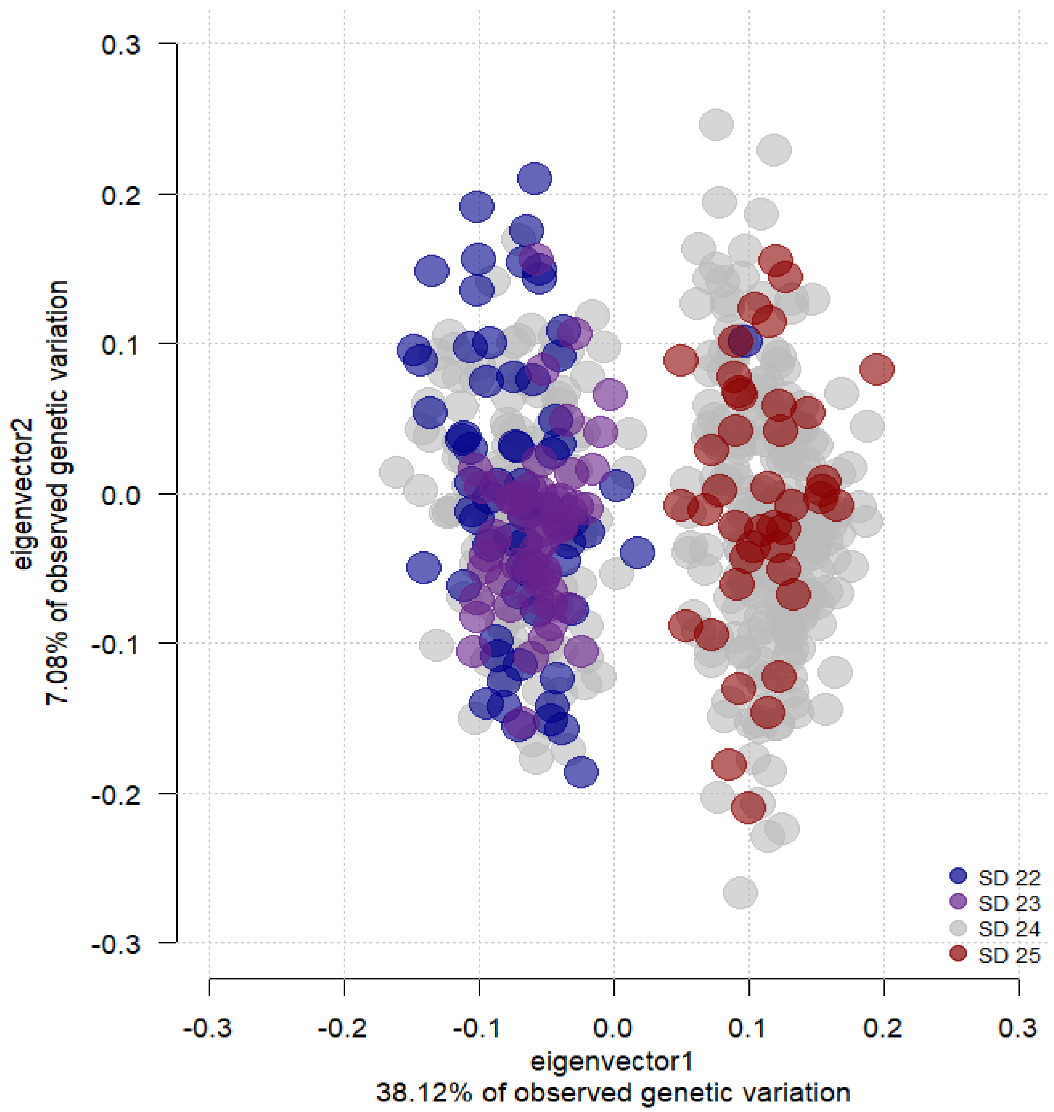


Figure 3: Principal Component Analysis of Baltic cod genotypes using the minimum SNP panel comprising 20 SNPs. Eigenvectors were inferred by using samples of spawning cod from ICES subdivisions (SD) 22 and 25. Points are color-coded according to locations within subdivisions. Total N=519. Note the blue point (one eastern Baltic cod caught in SD 22) in the eastern Baltic cod cluster on the right.

Otolith shape analysis

Images of entire and clean sagittal otoliths from genetically validated cod (N=507) were taken with a stereo microscope equipped with a digital microscope camera. Shape analyses on otolith images were conducted using the ShapeR package (Libungan & Pálsson 2015) in R. The contour of the otoliths was transformed into 45 shape coefficients using the normalized elliptical Fourier technique; the first three coefficients were used for standardization of otoliths with regard to size, rotation and starting point (48-3=45) (Figure 4).
The classification of cod individuals to their stock of origin based on otolith shape coefficients was performed using a linear discriminant analysis (LDA) and leave-one-out cross validation applying the MASS package (Venables & Ripley 2002) in R. Prior probabilities for LDA were set at 0.5 for both classification groups. Classification success of individuals (self-assignment) was calculated by comparing the results with the genetic assignment. Correct classification of this approach is presently 83%.

Otoliths with unknown stock origin were imaged and shapes were analysed following the procedure described above. Using the genetically validated baseline with stock-specific shapes, samples with unknown origin were assigned either to the WBC or EBC stock.

An extended baseline including juvenile samples from Denmark covering fish lengths from 18 to 37cm (N= 88) is presently tested.

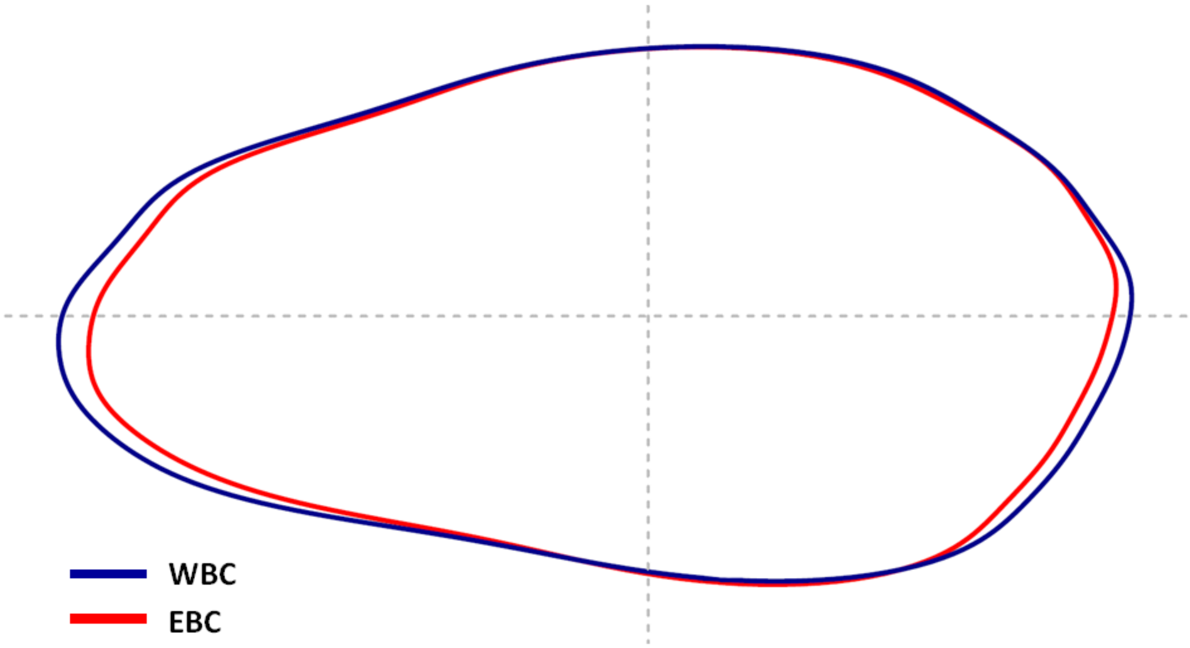


Figure 4: Mean otolith shape of the western Baltic cod and eastern Baltic cod stock based on the first 45 elliptical Fourier coefficients.

Results

Survey data

A time series of annual mixing proportions of WBC and EBC in SD 24 was developed using cod otoliths (N=17 206) from the German BITS in quarter 4 between 1977 and 2017 (Figure 5). In the late 1970s and in the 80s the mixing proportions of the cod stocks were relatively stable with an average mixing proportion of 42% WBC and 58% EBC in SD 24. In the first half of the 1990s the proportion of WBC increased to 68% in 1996 and decreased until the late 2000s to 27%. During the past ten years the mixing proportions of WBC and EBC in SD 24 were relatively stable with an average mixing proportion of 34% WBC and 66% EBC. The annual estimates from the German BITS deviate from the mixing proportions based on Danish commercial samples from 13 data years covering the period 1996 to 2017 by 5 to 30% for area 1 (12°-13°) and by 3 to 27% for area 2 (13°-15°).

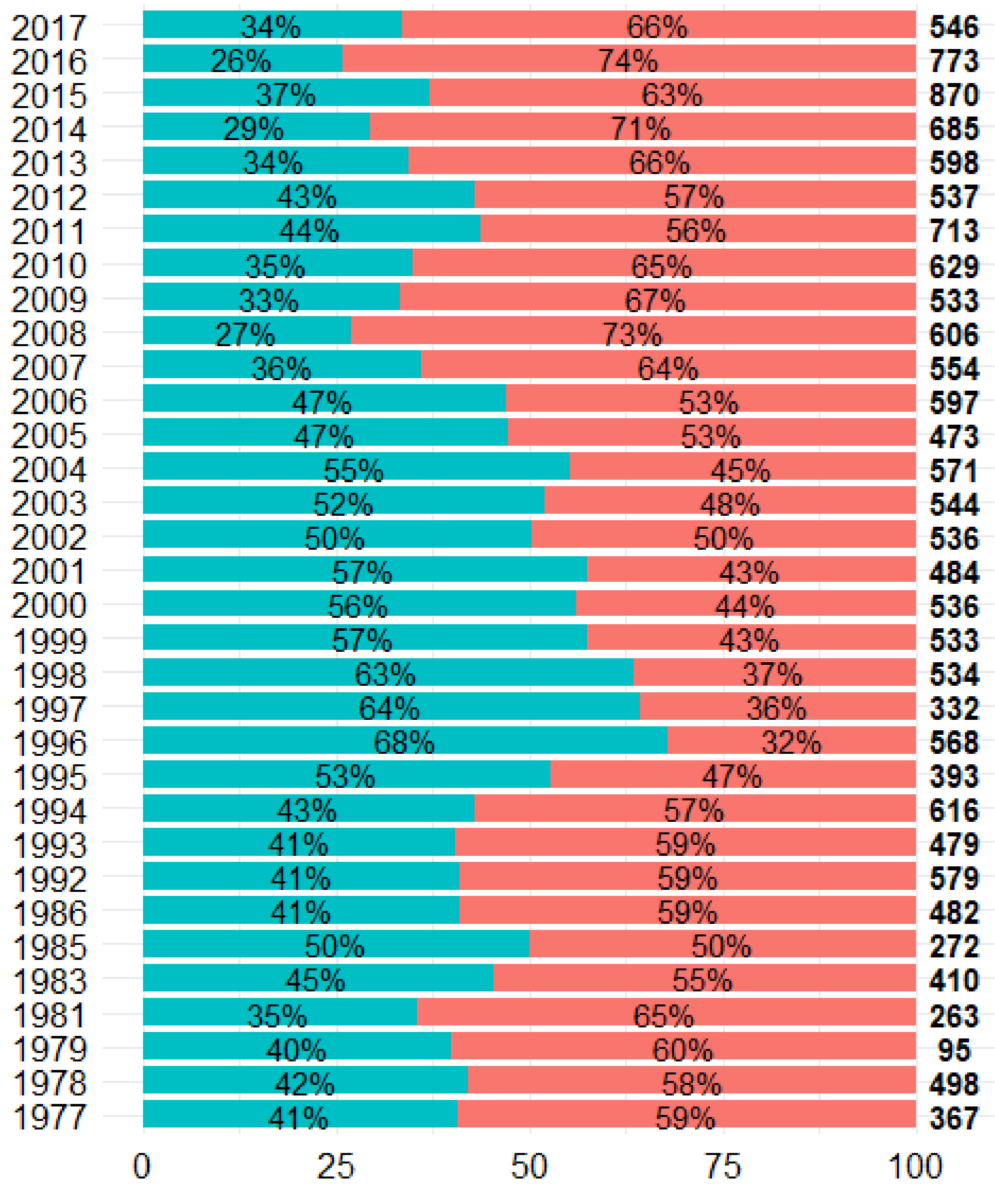


Figure 5: Mixing proportions of WBC (green) and EBC (red) in SD24 between 1977 and 2017 based on otoliths from the German BITS in quarter 4. Numbers of otoliths per year used in the shape analysis are presented on the right.

In addition to the overall mixing proportion in SD 24, the time series separated by longitude revealed a remarkable occurrence of WBC in the eastern part (14°-15°E) of the Arkona Basin (14 to 74%) and EBC in the western part (12°-13°E) of SD 24 (15 to 66%), suggesting that a mixing of the cod stocks may also occur beyond SD 24.

The comparison of mixing proportions based on cod samples from German BITS in quarter 4 and quarter 1 (selected years between 1995-2016 , N=3858 otoliths) did not show significant differences (1 to 10% deviation between quarters within the same year). This suggests that quarter 1 data can also be used as proxies for the mixing proportions of the quarter 4 survey within the same year, when otoliths from quarter 4 are missing. This mainly affects historical survey years in the 1980s.

To validate the classification success of the otolith shape baseline for historical samples, DNA from tissue attached to otoliths from selected years between 1979 to 1989 was extracted and successfully genotyped (N=155) using our minimal marker set of 20 SNPs. Classification success was calculated by comparing the genetic assignment with the assignment achieved from the otolith shape analysis. For the historical samples, 80% of the individuals were correctly assigned to their original stock.

Commercial data

Otolith shape analysis of German commercial samples (N=2864) from 2015 and 2016 detected considerably different mixing proportions between active and passive gear fisheries in SD 24. Lower proportions of WBC were found in catches using active gears (37% in 2015, 19% in 2016), confirming the results from BITS quarter 4 survey catches (37% in 2015, 26% in 2016), and higher proportions of WBC were revealed in catches using passive gears (64% in 2015, 66% in 2016), mainly fishing in water shallower than 20 m.

Additional analysis on German commercial cod samples originating from active gear fisheries in 2015 (N=2565) from SD 22, SD 25 and SD 26 showed substantial proportions of WBC in SD 25 and SD 26 (on average 16%), as well as EBC in SD 22 (18%). Unlike SD 24, differences in mixing proportions between active and passive gear fisheries were rather low in SD 22 (active: 82% WBC, passive: 84% WBC).

In addition, commercial samples from 2014 to 2016 provided by Sweden (N=1824; SD 24 and SD 25) and by Poland (N=738; SD 25 and SD 26) were investigated (passive gear fisheries only). Otolith shape analysis revealed a significant occurrence of WBC in the northern part of SD 24 and the western part of SD 25 (23 to 25%), as well as in the southern part of SD 25 and SD 26 (20 to 39%). Otoliths from SD 26 and 28 kindly provided by Latvia are presently analysed.

Conclusion

The comprehensive time series of mixing proportions of WBC and EBC based on German survey data from SD 24 is of great importance for establishing reliable reference points for mixed stocks with historically fluctuating stock levels. Moreover, otolith shape analysis of German commercial catches revealed differences in mixing proportions in cod from active (lower proportion of WBC) and passive gear fisheries (higher proportion of WBC), with the highest difference in SD 24. WBC seem to be distributed more in shallower coastal areas while EBC seem to dominate the deeper areas. Annual mixing proportions should consider samples from active and passive gear catches. Additionally, substantial mixing was detected also beyond SD 24, with considerable proportions of EBC in SD 22 and WBC in SD 25 and SD 26. These findings challenge some prevailing paradigms and verification by genetic analyses is needed.

References

Libungan LA, Palsson S (2015) ShapeR: An R Package to Study Otolith Shape Variation among Fish Populations. Plos One 10.

Patterson N, Price AL, Reich D, Reich D, Daly M (2006): Population structure and eigenanalysis. Plos Genetics, 2: 2074–2093.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004): GENECLASS2: A Software for Genetic Assignment and First-Generation Migrant Detection. Journal of Heredity, 95: 536–539.

Rannala B, Mountain JL (1997): Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences, 94: 9197–9201.

Tomkiewicz J, Tybjerg L, Holm N, Hansen A, Broberg C, Hansen E (2002): Manual to determine gonadal maturity of Baltic cod. DFU-rapport nr. 116-02, 52pp.

Venables WN, Ripley BD (2002): Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

WD 11 - Note on a method for estimating stock mixing proportions by otolith shape

Christoffer Moesgaard Albertsen[1,*]

[*]cmoe@aqua.dtu.dk
[1]National Institute of Aquatic Resources, Technical University of Denmark, Kemitorvet 201, DK-2800 Kgs. Lyngby, Denmark

ABSTRACT

A coherent model for estimating stock mixing proportions by otolith shape is introduced. The model can account for effects of fish length and year differences while providing unbiased stock mixing proportion estimates. The model is compared to a linear discriminant analysis through simulation studies.

1 Introduction

Otolith growth depends on species, stock, environmental factors, and individual differences. Therefore, the otoliths shape can be used to discriminate between mixing stocks such as the Eastern and Western Baltic cod. It is, however, important to account for other dominant factors on otolith shape such as fish length and year classes.

Closed contours can be described by normalized elliptical Fourier descriptors[1,2]. Therefore, Fourier descriptors have been used for analysing otolith shape (e.g.[3,4]). The normalized Fourier descriptors are subsequently used as features in a Linear Discriminant Analysis (LDA[5]). Based on a sample of fish with a known stock origin, LDA can be used to classify individuals with unknown stock origin to one of the stocks under consideration.

Classification by LDA is equivalent to a maximum likelihood procedure where the observed features, X_i, for an individual, i, are assumed to follow a multivariate normal where the mean depends on the stock, S_i, and individuals are classified by a maximum a posteriori probability rule. That is,

$$X_i \mid S_i = s \sim \mathcal{N}(\mu_s, \Sigma).$$

When a vector of prior probabilities of belonging to each stock is given, $\check{\pi}$, the posterior probability that individual i belongs to a stock given the observed features can be calculated by Bayes' theorem,

$$P(S_i = s \mid X_i) = \frac{f(X_i \mid S_i = s) \cdot \check{\pi}_s}{\sum_{s=1}^{N_S} f(X_i \mid S_i = s)},$$

where f is the density of $X \mid S_i$. An individual is classified, or allocated, to a specific stock (C_i) by the rule,

$$C_i = \operatorname{argmax}_s P(S_i = s \mid X_i)$$

We denote the resulting stock proportions by $\tilde{\pi}$. That is,

$$\tilde{\pi}_s = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{C_i = s}(i).$$

It is well known that estimating stock mixing proportions by the numbers classified to each stock provides biased results (e.g.,[6–9]; See Figure 1 for an illustrated example). The bias will depend on both the misclassification probabilities and the true stock mixing proportion, π. Defining the confusion matrix Π such that the element Π_{nm} is the probability of classifying an individual from stock m to stock n, the expected estimated mixing proportion can be expressed by

$$E(\tilde{\pi}) = \Pi\pi, \tag{1}$$

which, in general, does not equal π. Naturally, Π depends on the classification procedure.

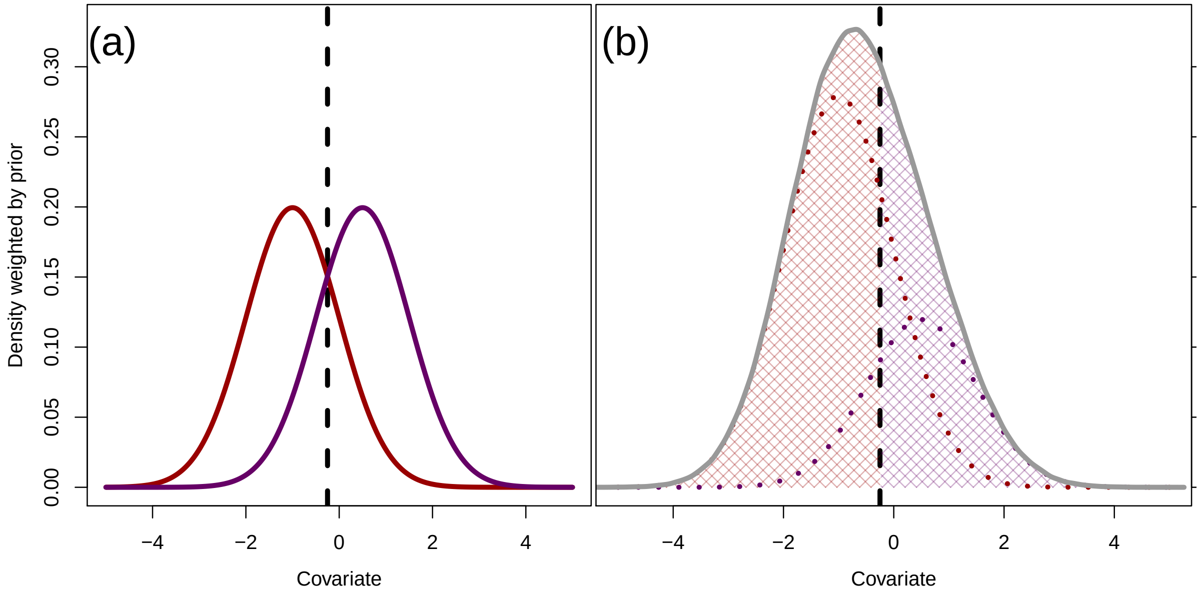


Figure 1. Illustration of classification by Linear Discriminant analysis. In panel (a), the red and purple full lines are the densities of the covariate estimated from individuals with a known stocks and prior probabilities of 0.5. The black dashed line is the resulting classification boundary. To the left of the black line, individuals are classified to the red stock. In panel (b), the grey full line illustrates the density of individuals with an unknown stock origin (based on 70% red and 30% purple, illustrated by the corresponding dotted lines). Based on the data with known stock origin in panel (a), 61% are classified to the red stock (red shaded area), while 39% are classified to the purple stock (purple shaded area).

While the LDA classifier provides biased stock mixing proportions if $\tilde{\pi}$ is used, equation (1) can be used to correct the bias. When Π is known without error, the mixing proportions can be corrected by using $\Pi^{-1}\tilde{\pi}$. In this case,

$$E(\Pi^{-1}\tilde{\pi}) = \Pi^{-1}E(\tilde{\pi}) = \Pi^{-1}\Pi\pi = \pi.$$

This procedure has previously been used for bias-correcting estimated mixing proportions of commercial catch of Eastern and Western Baltic cod[4]. The bias-correction method does not provide confidence intervals; however, approximate uncertainties can be calculated from the multinomial distribution, assuming the parameters of the Gaussian distributions and Π are known without error. Alternatively, bootstrap methods can be used[9].

The linear discriminant analysis can not include covariates such as fish length or year of capture to correct the mean estimates. This has been circumvented by a stepwise procedure[4]. In the stepwise procedure, a regression is first made for each feature on fish length. Afterwards, the data is split into length groups for which LDAs are made independently.

This note describes a coherent model for analysing otolith shape and classifying individuals while accounting for effects of fish length and yearly differences within the model. For fish with a known stock origin, the model extends the LDA to include length and year effects. For fish with an unknown stock origin, the model provides unbiased mixing proportion estimates. Unbiased mixing proportion estimates are found by realising that data with unknown stock origin comes from a mixture of the normal distributions estimated from data with known stock origin.

2 A coherent model for estimating mixing proportions

The coherent model is an extension of the linear discriminant analysis. For a given stock (S_i) and covariates (the row vector Z_i), a set of observed features (X_i) is assumed to follow a multivariate normal,

$$X_i \mid Z_i, S_i = s \sim \mathcal{N}(Z_i\beta_s, \Sigma).$$

with density $f(X_i \mid Z_i, S_i = s)$. Here, β_s is a matrix of parameters. The number of columns corresponds to the number of features, while the number of rows corresponds to the number of covariates. In the simplest case where Z_i is 1, the model reduces to an LDA. Note that the model can be further generalized by letting the covariance depend on the stock; however, this will drastically increase the number of parameters. The formulas below would remain the same.

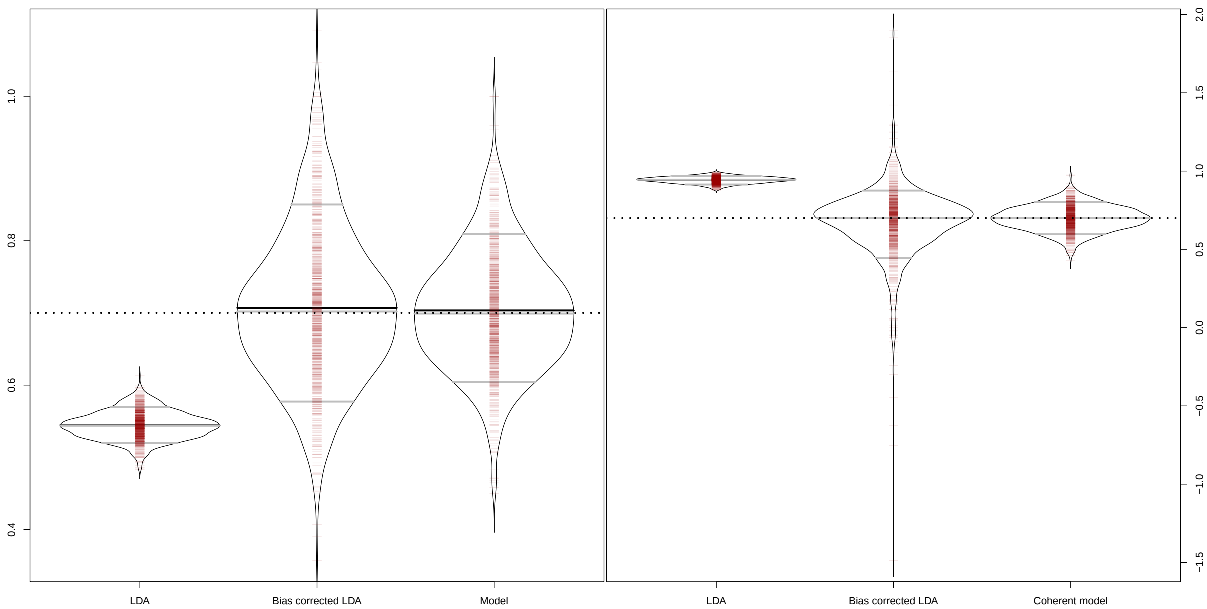


Figure 2. Bean plots of estimated stock mixing proportion for the first stock. In the top panel, the training data is balanced. In the bottom panel, the training data has the true stock mixing proportions. The black dotted line shows the true value. Grey lines in the beans show the 10%, 50%, and 90% quantiles, while the black lines show the mean. Short red lines show individual data points.

The model differs further from the LDA for individuals with unknown stock origin. For data where S_i is unknown, the features are assumed to follow a mixture of the possible stocks,

$$f(X_i \mid Z_i) = \sum_{s=1}^{N_s} f(X_i \mid Z_i, S_i = s)\pi_s.$$

The new stock mixing parameters can be made dependent on covariates like capture area, year, or sex.

Again, posterior probabilities are calculated by

$$P(S_i = s \mid X_i, Z_i) = \frac{f(X_i \mid Z_i, S_i = s) \cdot \hat{\pi}_s}{\sum_{s=1}^{N_S} f(X_i \mid Z_i, S_i = s)},$$

where $\hat{\pi}$ is the maximum likelihood estimate of π. An individual is still classified to a specific stock by the maximum a posteriori probability,

$$C_i = \text{argmax}_s P(S_i = s \mid X_i, Z_i).$$

Note that while the posterior probabilities are consistent with the unbiased model, the number allocated to each stock need not correspond to estimated stock mixing proportions for a finite sample.

3 Simple simulation study

To illustrate the applicability of the model, a simulation study is conducted. The simulation study considers a simple case where the LDA is the true model, and there are two stocks. The group means are 0.5228008 and 0.5376781, respectively, while the variance is 0.0007151752124176. These values correspond to the D1 normalized elliptical Fourier descriptor of a 50 cm fish from the Danish baseline data (Figure 4). The study consists of 1,000 replications of two cases. In each case, 1,000 samples with a known stock origin, and 1,000 samples with an unknown stock origin were simulated. In both cases, the samples unknown stock origin are simulated such that 70% belong to the first stock. In the first case, the samples with known stock origin are simulated with 50% from each stock. In the second case, the same proportions are used for the samples with known and unknown stock.

For each simulated data set, an LDA is fitted on the data with a known stock origin, the samples with an unknown stock origin are classified, and the stock mixing proportions, $\tilde{\pi}$, are calculated. Further, bias-corrected estimates, $\Pi^{-1}\tilde{\pi}$, are calculated

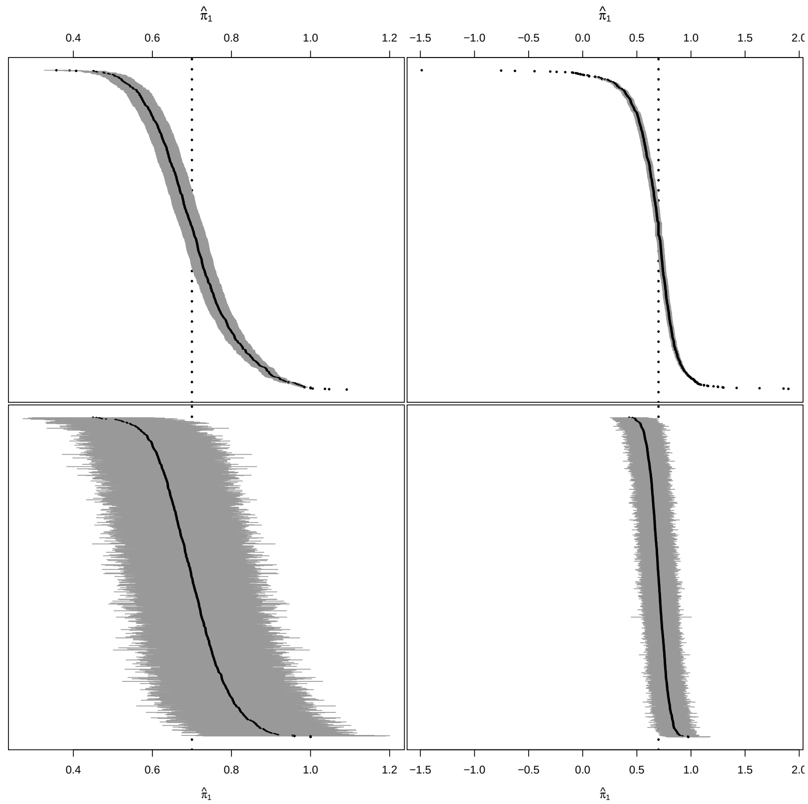


Figure 3. Coverage plot for the estimated stock mixing proportions. Each black dot corresponds to an estimated mixing proportion. The grey lines are the corresponding Wald confidence intervals for the bias-corrected LDA based on the binomial distribution (top panels) and the coherent model (bottom panels). In the left panels, the training data is balanced. In the right panels, the training data has the true stock mixing proportions. Bias-corrected estimates with an absolute value above 2 have been removed.

and their standard errors are approximated from a binomial, $\sqrt{p(1-p)/N}$. Π is estimated by leave-one-out cross-validation on the data with known stock origin. Finally, the coherent model is used to obtain maximum likelihood estimates of the stock mixing proportions along with their standard errors. Confidence intervals are constructed on the proportion scale. Therefore, they are not restricted to be between 0 and 1. For applications, confidence intervals can be calculated on the scale the parameter is estimated on, and transformed to the proportion scale, thereby restricting them to be between 0 and 1.

From the simulations, it is evident that the naive stock mixing proportions from the LDA are biased (Figure 2). When a balanced data set is used to estimate the parameters of the LDA, and as prior for the classifications, the average estimated stock mixing proportion is 0.544. The true value is 0.7. Both the bias-corrected values and the model estimates provide the true stock mixing proportion on average. Even when the true stock mixing proportion is used as prior, the naive stock mixing proportions from the LDA are biased. In this case, the average stock mixing proportion is 0.943. Again, both the bias-correction and the coherent model provides the correct mixing proportions. In both cases, the variability of the model estimate is lower than the bias-corrected values. Further, the model estimate is inherently restricted to be between 0 and 1. It is possible to modify the bias-correction procedure to have the same constraint; however, this is outside the scope of this note.

An advantage of the coherent model is seen when the coverage of confidence intervals are considered. When a step-wise procedure is used, such as the bias-corrected LDA, it can be difficult to properly accumulate the uncertainties from each step. This is accounted for in the coherent model. In the case where a 50% prior is used, the coverage of the bias-corrected 95% confidence intervals is 0.237 (Figure 3). In contrast, the coverage of the confidence intervals provided by the model is 0.966. Likewise, the coverages when the true stock mixing proportions are used as priors are 0.173 and 0.959, respectively.

4 Otolith shape of Baltic cod

Unfortunately, the world is not as simple as the simulation study. In real cod otoliths, the normalized elliptical Fourier descriptors depend on fish length, and may even differ from year to year. Figure 4 illustrates the effect of fish length on the

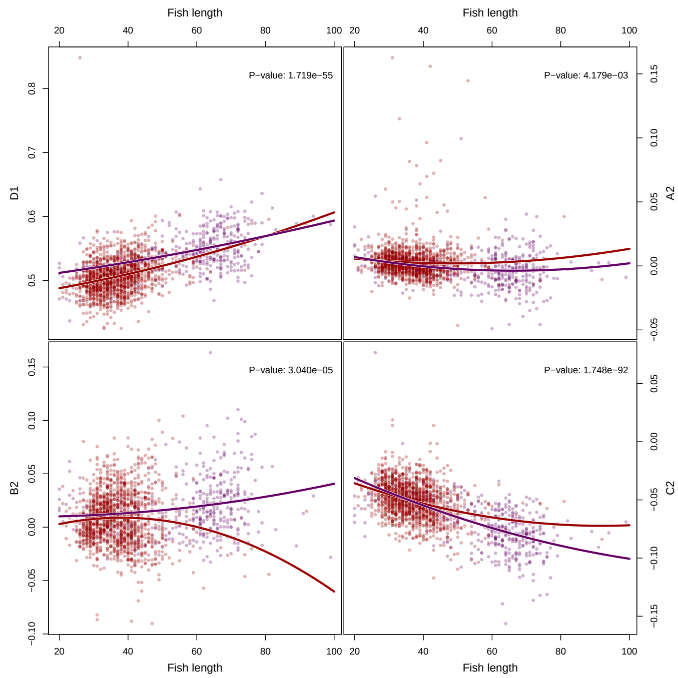


Figure 4. Four normalized elliptical Fourier descriptors for Danish otolith data with a known stock origin as a function of fish length. Only left otoliths are used. Full lines are fitted second-degree polynomials for Eastern (red) and Western (purple) Baltic cod. Each point represents the value of a single otolith. The indicated P-values are based on individual likelihood ratio tests of the hypotheses of no length effects.

normalized elliptical Fourier descriptors D1, A2, B2, and C2. Not only is there a length effect on the descriptors, but the length effect depends on the stock.

While many Fourier descriptors are needed to accurately describe an individual otolith contour, most of the stock difference can be described by the first few harmonics (Figure 5). Fitting the stock-wise second degree polynomial on length for the first 200 harmonics (800 descriptors), the estimated stock difference parameters can be extracted. The model includes a difference in the intercept, the linear effect, and the quadratic effect. For the first few harmonics, the estimated stock difference is large, while it decreases to zero as the harmonic number grows. This means that only the first harmonics are important for discriminating stocks, while higher harmonics determine the specific shape of an individual otolith.

5 Importance of accounting for length effects

To investigate the importance of accounting for length effects, a simulation study is conducted. In the simulation study, data from the Danish otoliths from 2011 and 2012 are used as input data with known stock origin. Only the D1 descriptor is considered for simplicity.

Data with unknown stock origin is simulated from the stock-wise polynomial fit shown in Figure 4. Three different cases for simulating length effects are considered. In the first case, lengths are simulated uniformly. In the second case, lengths are simulated from a normal distribution. The Eastern stock is simulated with mean 35 and standard deviation 10. The Western stock is simulated with mean 65 and standard deviation 20. this is similar to the length distribution of the data with a known stock origin. In the third case, the Eastern and Western normals are switched. Each of the three cases is combined with Eastern mixing proportions of 0.8, 0.5, and 0.3. The data sets contained 1000 observations of unknown stock origin.

In all nine cases, the coherent model provides reasonable stock mixing proportion estimates (Figure 6), while the LDA based estimates are only close to the true value in two cases. The LDA provides reasonable estimates when the true mixing proportion is 0.5, and the length distribution is similar to the data with known stock origin. The LDA also provides reasonable estimates when the length distributions of the Eastern and Western stocks are switched. In this case, the average classification success of the unknown samples are, however, only 32.5%, but because the mixing proportion is 0.5, switching the stocks does not matter for the estimate.

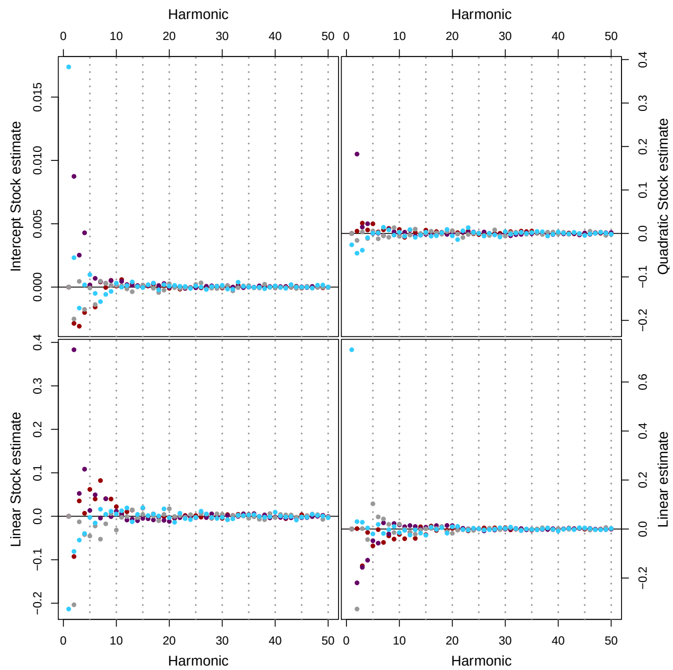


Figure 5. Estimated parameter values from a stock-wise polynomial fit as a function of elliptical Fourier harmonic. For each harmonic there are four descriptors: A (red), B (purple), C (grey), and D (blue).

6 Accounting for year effects

The final issue investigated is the presence of yearly differences in the Fourier descriptors. A year effect shifts the values of mean parameters as illustrated in Figure 7. In an LDA, the estimation procedure will be unaware of such effects. Consequently, data with unknown stock origin will be classified based on data with different mean values, leading to a wrong result.

In the coherent model, Z_i can be modified to account for yearly differences in the mean parameters. To illustrate, a single simulated data set is constructed. For each stock, 1,000 observations are simulated with known stock origin from the distributions shown in Panel (a) in Figure 7. Further, 1000 observations from each stock are simulated with unknown stock origin from Panel (b) in Figure 7.

Fitting the model without accounting for the year effect results in an estimated (standard error) stock mixing proportion of 0.228 (0.022), far from the true value of 0.5. This is caused by the fact that the estimated mean parameters are influenced by both the data with known and unknown stock origin since we assume they come from the same distributions. Fitting the model only to data with a known stock origin, the estimated stock-wise means are -0.956 (0.031) and 0.5 (0.031). When data with unknown stock origin are included to estimate the mixing proportion, the estimated stock-wise means become -0.941 (0.031) and 0.569 (0.028), respectively. That is, the model finds the a compromise between the two data generating distributions for each stock, but the data with known origin is weighed higher.

When a year effect is included, shifting both stock-means by the same value, the estimated stock mixing proportion becomes 0.535 (0.048), and the resulting 95% confidence interval includes the true value. For this model, the estimated mean parameters are -0.956 (0.031) and 0.5 (0.031) when only data with known stock origin are used. When data with unknown stock origin are included, the estimated means become -0.968 (0.029) and 0.511 (0.029), while the estimated year effect is 0.547 (0.547). Now, the model no longer needs to find a compromise, leaving the mean values for the data with known origin almost as they were, but has the freedom to estimate the shift in the values between years. Even when one of the years only had data with unknown stock origin. Naturally, the estimation will be aided by having data with known stock origin from every year, and in some cases, it may even be necessary.

7 Conclusion

In this note, a method was introduced that accounts for length and year differences while estimating stock mixing proportions in a single coherent model. Parameters, including the mixing proportions, are estimated by maximum likelihood. Therefore, reasonable stock mixing proportions could be estimated in each scenario investigated. Further, realistic quantifications of their uncertainties could be calculated. In particular, the confidence intervals calculated in the first simulation study had the correct

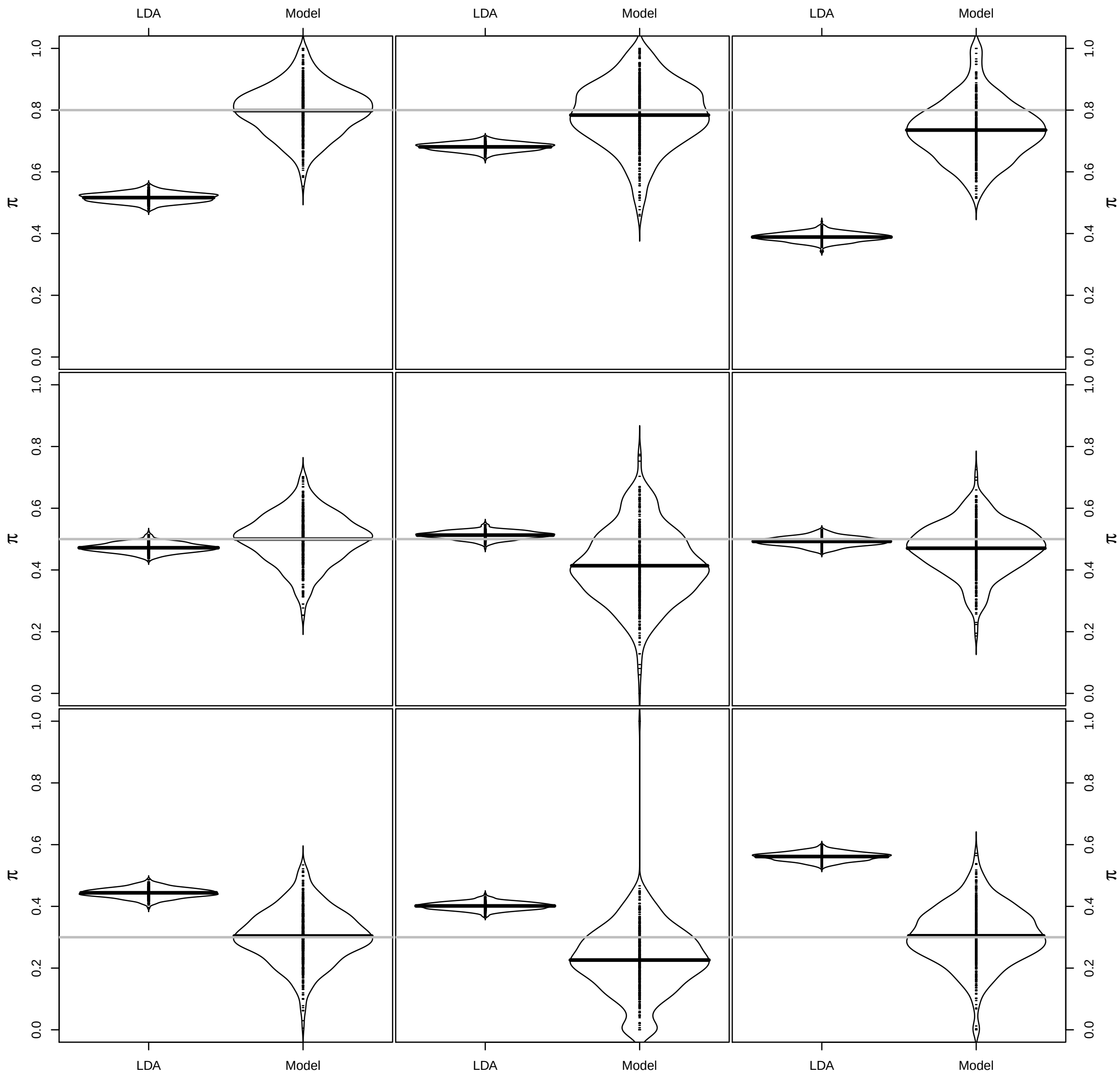


Figure 6. Estimated stock mixing proportions for the nine length-effect cases. Each row shows the result for different true mixing proportions. In the first column, data with unknown stock origin has a uniform length distribution. In the second column, data with unknown stock origin follow normal distribution where the Eastern stock is smaller - as the known data. In the second column, data with unknown stock origin follow normal distributions where the Western stock is smaller - unlike the known data.

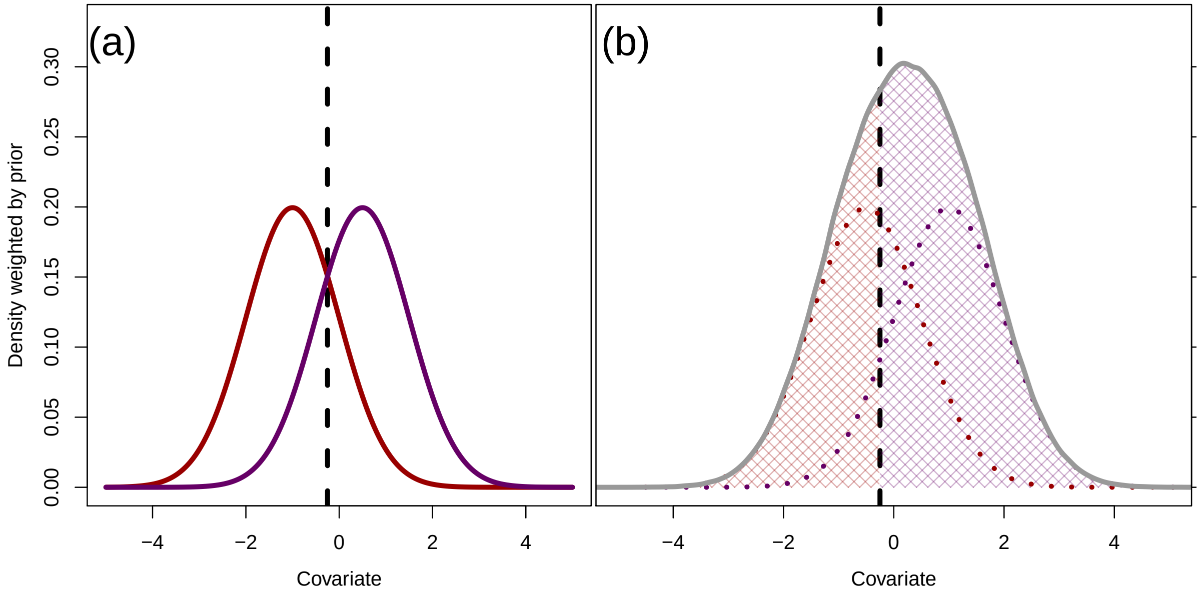


Figure 7. Illustration of year effects on the mean parameters in a Linear Discriminant analysis. In panel (a), the red and purple full lines are the densities of the covariate estimated from individuals with known stocks and prior probabilities of 0.5. The black dashed line is the resulting classification boundary. To the left of the black line, individuals are classified to the red stock. In panel (b), the grey full line illustrates the density of individuals with an unknown stock origin (based on 50% red and 50% purple, illustrated by the corresponding dotted lines). The stock means are shifted to the right, representing a year effect. Based on the data with known stock origin in panel (a), 35% are classified to the red stock (red shaded area), while 65% are classified to the purple stock (purple shaded area).

coverage. As a result, asymptotic normality of the estimator can be used to investigate the number of samples needed to obtain a certain standard error of the stock mixing proportions, and how the samples are best divided between known and unknown stock origin.

The model can be extended to include various effects; however, like any estimation procedure, overfitting is a concern. In a model of N features, the simple LDA model includes $2 \cdot N$ mean parameters, N variance parameters and $(N \cdot N - N)/2$ correlation parameters. A number that grows rapidly. To avoid overfitting, a feature selection procedure can be used, or a penalty can be put on the mean parameters to shrink the stock difference towards zero.

If the model specification used does not adequately describe the data at hand, the estimated stock mixing proportions will be wrong. The final simulation study outlined a check for the adequacy of the model. If the estimated mean parameters change substantially when data of unknown stock origin are included, the estimated stock mixing proportions are likely wrong. This can be corrected by including, for instance, length or year effects in the model. The model can be combined with survey index or stock assessment models to accurately accumulate uncertainty from the procedure.

References

1. Kuhl, F. P. & Giardina, C. R. Elliptic fourier features of a closed contour. *Comput. Graph. Image Process.* **18**, 236 – 258 (1982). DOI https://doi.org/10.1016/0146-664X(82)90034-X.

2. Ferson, S., Rohlf, F. J. & Koehn, R. K. Measuring shape variation of two-dimensional outlines. *Syst. Zool.* **34**, 59–68 (1985).

3. Campana, S. E. & Casselman, J. M. Stock discrimination using otolith shape analysis. *Can. J. Fish. Aquatic Sci.* **50**, 1062–1083 (1993). DOI 10.1139/f93-123.

4. Hüssy, K. *et al.* Evaluation of otolith shape as a tool for stock discrimination in marine fishes using baltic sea cod as a case study. *Fish. Res.* **174**, 210–218 (2016). DOI 10.1016/j.fishres.2015.10.010.

5. Fisher, R. A. The use of multiple measurements in taxonomic problems. *Annals Eugen.* **7**, 179–188 (1936). DOI 10.1111/j.1469-1809.1936.tb02137.x.

6. Healy, J. D. The effects of misclassification error on the estimation of several population proportions. *The Bell Syst. Tech. J.* **60**, 697–705 (1981). DOI 10.1002/j.1538-7305.1981.tb00257.x.

7. Chhikara, R. S. Error analysis of crop acreage estimation using satellite data. *Technometrics* **28**, 73–80 (1986).
8. Bélanger, J. & Gagnon, D. Classification and proportion estimation. *Aust. J. Stat.* **35**, 19–28 (1993). DOI 10.1111/j.1467-842X.1993.tb01309.x. `https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-842X.1993.tb01309.x`.
9. Wilkins, M. F., Boddy, L. & Morris, C. W. Proportion estimation with confidence limits. *J. Microbiol. Methods* **43**, 55–64 (2000). DOI 10.1016/S0167-7012(00)00204-9.

WD 12 - Issue list Western Baltic Cod

	Western Baltic cod stock and area (22-24)	Years	Activity	Country, Scientist(s)	Priority
	Input data				
1	Update the catch data time series	All yrs concerned	WD	DK, DE	done
4	Information on number of boxes per size sorting category (port sampling) (PSU)	At least last 5 yrs	WD	DK, SWE	1
5	Comparison of length, weight and age distributions in SD22 based on current data	At least last 5 yrs	WD	DK, DE	done
6	Sensitivity analysis of input data (biological sampling): (a) Pseudostock definitions SD 22 in IC (pseudostock 1: DK caton + DK biological data + DE caton; pseudostock 2: DK caton + DE caton + DE biological data) (b) Upload data for each pseudostock (IC) (c) Raising in IC (d) Comparison of results	At least 2 recent yrs	Upload, stock coordinat or	(a) ICES (b) DK, DE (c) DE (d) DE, DK	1
	Age reading				
1	Consider DE age validation results from SD22 in age reading routine (1,9 mm diameter of first ring); discuss progress in otolith preparation (broken vs sliced)	? yrs	training course or workshop with age readers	DK, SWE, DE	Have been solved
2	Organize yearly exchange of otoliths in order to include an age error matrix in the routine assessment (consider experience from otolith exchange in 2015)	2015 done, 2016 onwards?	Otolith exchange SD22	DE, DK SWE	Partly done
	Age errow matrix			DE (DK, SWE)	
	Mixing SD24				
1	Why restrict mixing to SD24? Mixing between SD22-24-25?	Since the 1980s	WD	DE	3/ has been tested but presently not enough data
2	Otolith shape: Extending and completing the existing time series	See WKBALTCOD 2015 report	National labs	DE, DK	More years have been included
3	Otolith shape: More years with genetic validation	See WKBALTCOD 2015 report	National labs	DK, DE	More tears with genetics

					have been included
4	Otolith shape: Compare data from same years	2015	exchanging otolith shape data	DK, DE	Inter-sessional workshop conducted in July 2018
5	Documentation of present and historical splitting procedure (how many? Where? When? Biological data? Genetic approach/documentation? Source of samples? etc)	-	WD	DK, DE	WD at the benchmark
6	Otolith shape: Organization of future otolith sampling and analysis to achieve an improved spatial and temporal coverage	2016 onwards	Establish sampling scheme	DK, DE	Planned for 2019
	Have separate mixing in the passive and active gears, length				Partly conducted
7	Assign catches to area 1 and area 2 (SD24) (or more)	1995+	Report to stock assessor	DK, DE, SWE, LAT, POL	done
8	Include the split in the survey data				Been tested
	Survey				
1	BITS in SD24 has to be included in both WBC and EBC assessment	? yrs	Establish method	DK, DE, WGBIFS	Been tested
2	Improvement of BITS design (e.g. additional samples SD22?)	2017+	Discussion	DK, DE, WGBIFS	ongoing
	Comparative fishery of SOLEA and new Havfisken in SD24	Q1 2017 (BITS)	WD	DK, DE	done
	Stock weight from survey - caspers model with smooth ALK – check time series	Time series			done
	Future data				
1	Pound net samples DE SD22: recruitment index data	2011 onwards	WD	DE	done
	Include more fleets in IC				
	(gilnetters/ pelagiv trawl/ bottom trawl> 105/ bottom trawl<105/ longliner				Not conducted
	Reference points				
2	To check if they need to be revised according to new guidelines				done
3	**Recreational data**				
	Include Danish and Swedish recreational data				done
	Use biology from recreational surveys on German data				Partly

	Model fit				
	Use the latest year to calculate the short term – not the year -1				done
	Random walk in recruitment – change to BH or R				Tested but not adopted
	Strong RETRO pattern in SSB	No problem			
	Strong RETRO pattern in the catch				
	De couple the catchabilities for age 4 and 5 in survey				done
	Use a new F bar –different ages weighted by importance in fisheries.	Maybe to strange			Not conducte d
	Seperate fleet in the model (recreational fisheries and commercial catch)			DK	tested
	Maturity ogive			DK / GE	updated

WD 13 - WESTERN BALTIC COD CATCH DATA

Margit Eero, Marie Storr-Paulsen, DTU Aqua

1. APPLYING STOCK SPLIT ON COMMERCIAL COD CATCHES IN SD 24

Data on relative proportions of Eastern and Western Baltic cod in SD24

In previous stock assessments (WGBFAS 2018), stock splitting information has only been available from Danish samples. The methodology used to identify relative proportions of EB and WB cod in Danish commercial catches is described in Hüssy et al (2016 a and b). Stock splitting proportions are calculated separately for sub-areas 1 and 2 (Fig. 1), due to east-west gradient in stock mixing proportions (Hüssy et al. 2016b). The stock splitting proportions in Danish data are available from 1996 onwards, with several years of gaps in the time series (Fig. 1).

For the benchmark data meeting in 2018, proportions of EB and WB cod in German commercial catches in SD24 were provided for some later years (Fig. 2). Only data from Active gears were used (2005, 2010, 2015-2016). For the historical period (1977-1995), proportions of EB and WB cod were made available from German historical survey (1977-1986), supplemented by stock proportions derived from BITS survey (1992-1995). These stock proportions from surveys use only the cod above 30cm in length. The stock proportions in German data (survey and German commercial catch) are derived using the methodology described in WD_EBC_WBC1.

For a combined time series of stock proportions, DK and DE mixing proportions were combined. For the years, where stock split from both countries was available (2005, 2010, 2015-2016), these were averaged, weighted by landings of DK and DE (Active gears), respectively. For years where data on stock proportion were not available, extrapolations (averages of adjacent years) were applied (Fig. 3).

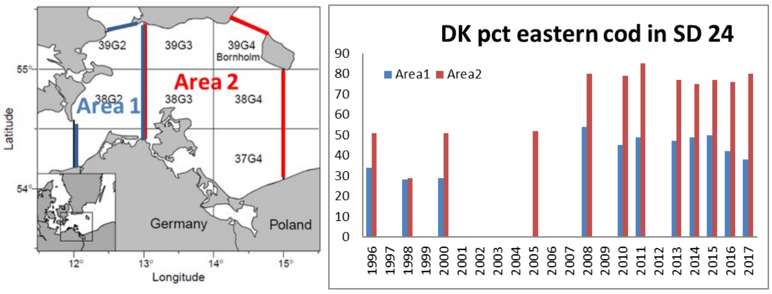


Fig. 1. Left panel: sub-areas used to apply stock split for cod catches in SD 24. Right panel: Proportions of EB cod in sub-areas 1 and 2, in years for which data are available from Danish commercial samples.

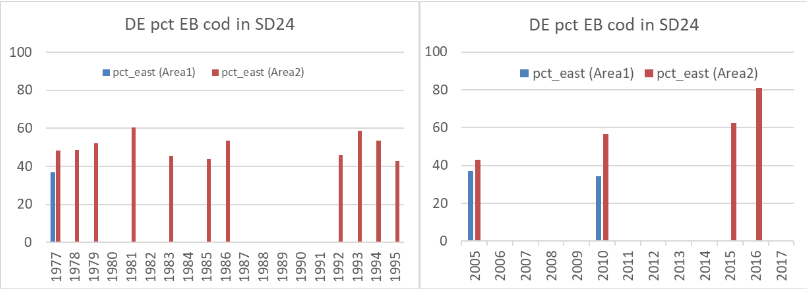


Fig. 2. Proportions of EB cod in sub-areas 1 and 2, from German survey (left panel) and from German commercial samples from active gears (right panel).

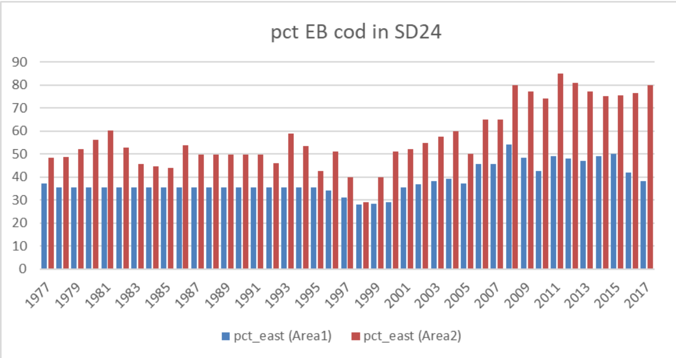


Fig 3. Proportion of EB cod in SD 24, by sub-areas, including extrapolation for years with missing data.

Separating total cod landings in SD24 to stocks

For each country, relative proportion of cod landings in sub-areas 1 and 2 within SD24 were derived from national data. For earlier years, where this information was not available, extrapolations of the landings distribution from more recent years were applied (Fig. 4).

For DK, the landings in SD 24 from 1996 onwards were split using DK stock proportions, separately by sub-areas. For example, the EB cod landings in sub-area 1 in a given year (y) were derived:

$$Catch\_EB\_Area1_y = Catch\_SD24_y * Prop\_Catch\_Area1_y * Prop\_EBcod\_Area1_y$$

where Catch_SD24 is total DK cod catch in SD 24 in a given year; Prop_Catch_Area1 is the proportion of DK cod catch in Area1, and Prop_EBcod_Area1 is the proportion of EB cod in Area1.

For years and sub-areas, where DE commercial catch split data were available (2005, 2010, 2015-2016), these data were applied to distribute DE commercial Active gear catches to stocks, in a similar way as for DK data.

To distribute the cod landings to stocks in other years and for other countries (OTHER), first the combined proportion of international landings in sub-areas 1 and 2 was derived. This was calculated as an average for DK, DE, SWE and POL, weighted by the total landings of these countries in SD24. Combined stock split using all available information was applied (Fig. 3), separately by sub-areas.

These steps resulted in stock specific landings for DK, DE and OTHER, by sub-areas 1 and 2, which where summed up to obtain total landings of EB and WB cod in SD 24 (Fig. 5).

For EB cod assessment, the split of landings in SD24 to stocks was extended further back to 1965. This was done applying average proportions of landings in sub-areas 1 and 2, and average stock mixing proportions from 1977-1979.

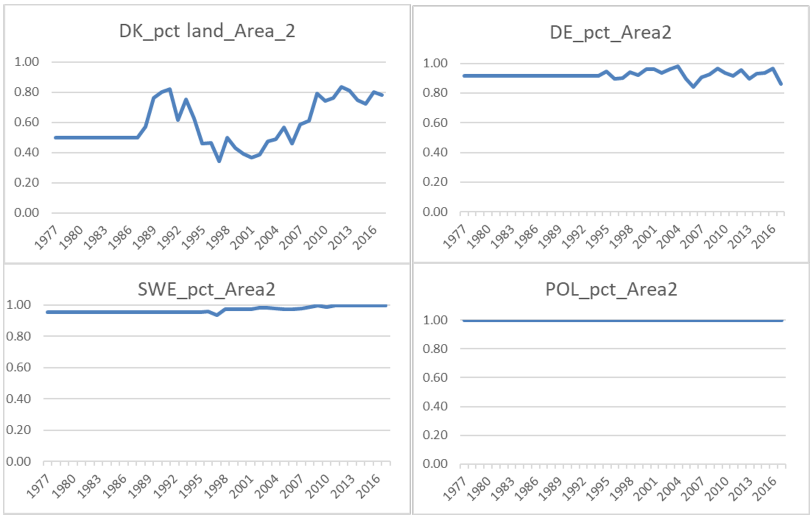


Fig. 4. Proportion on cod landings in sub-area 2 within SD 24, by country.

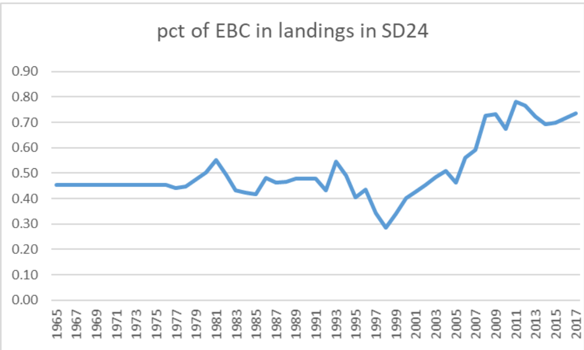


Fig. 5 Proportion of EBC in total cod landings in SD 24.

Separating total cod discards in SD24 to stocks

Cod discards in SD 24 are allocated to stocks from 1994 onwards, i.e. the time series of stock assessment for WB cod in previous assessments (WGBFAS 2018). The total estimated discards in tons in SD24 in 1994-2017 were allocated to stocks using annual average stock mixing proportions. These were derived from averaging stock splitting keys in sub-areas 1 and 2, weighted by proportion of landings in these subareas, by years. The resulting proportion of EB and WB cod in SD 24 was multiplied with total cod discards in SD24, to obtain discards for EB and WB stock in SD 24.

Age/size structure of commercial catch in SD 24

In previous assessments (WGBFAS 2018), data on age or size structure of cod catches in SD24 have not been used. Instead, the catch at age in SD22 has been raised to account for tons on WB catches in SD 24. Similarly, catch at length in SD25 has been raised to account for EB cod catch in SD24. This assumes that the WB cod in SD 24 have the same age structure as the WB cod in SD 22, and the EB cod in SD 24 have the same size structure as the EB cod in SD 25.

At benchmark data meeting 2018, this assumption was evaluated using data on observed length distributions in SD 24 in 2000-2013 (compiled in Intercatch). This was done by comparing the calculated catch at length in SD24 with the observed. The calculated catch at length was obtained by summing up i) the WB cod fraction of catch at length in SD 24, obtained when raising catch a length in SD22 by the WB cod tons from SD 24; and ii) the EB cod fraction of catch at length in SD24, obtained by raising catch a length in SD25 by the EB cod tons from SD 24.

The calculated and observed catch at length in SD24 were relatively similar (Fig.6), indicating that the assumption used in the assessment is reasonable. Therefore, the approach previously applied in the assessment was suggested to be continued. Alternatively, observed catch at length in SD24 could be split between stocks, similarly to tons. However, it is presently unclear how the proportion of stock mixing might change with the length of the fish, depending on year-class strength etc.

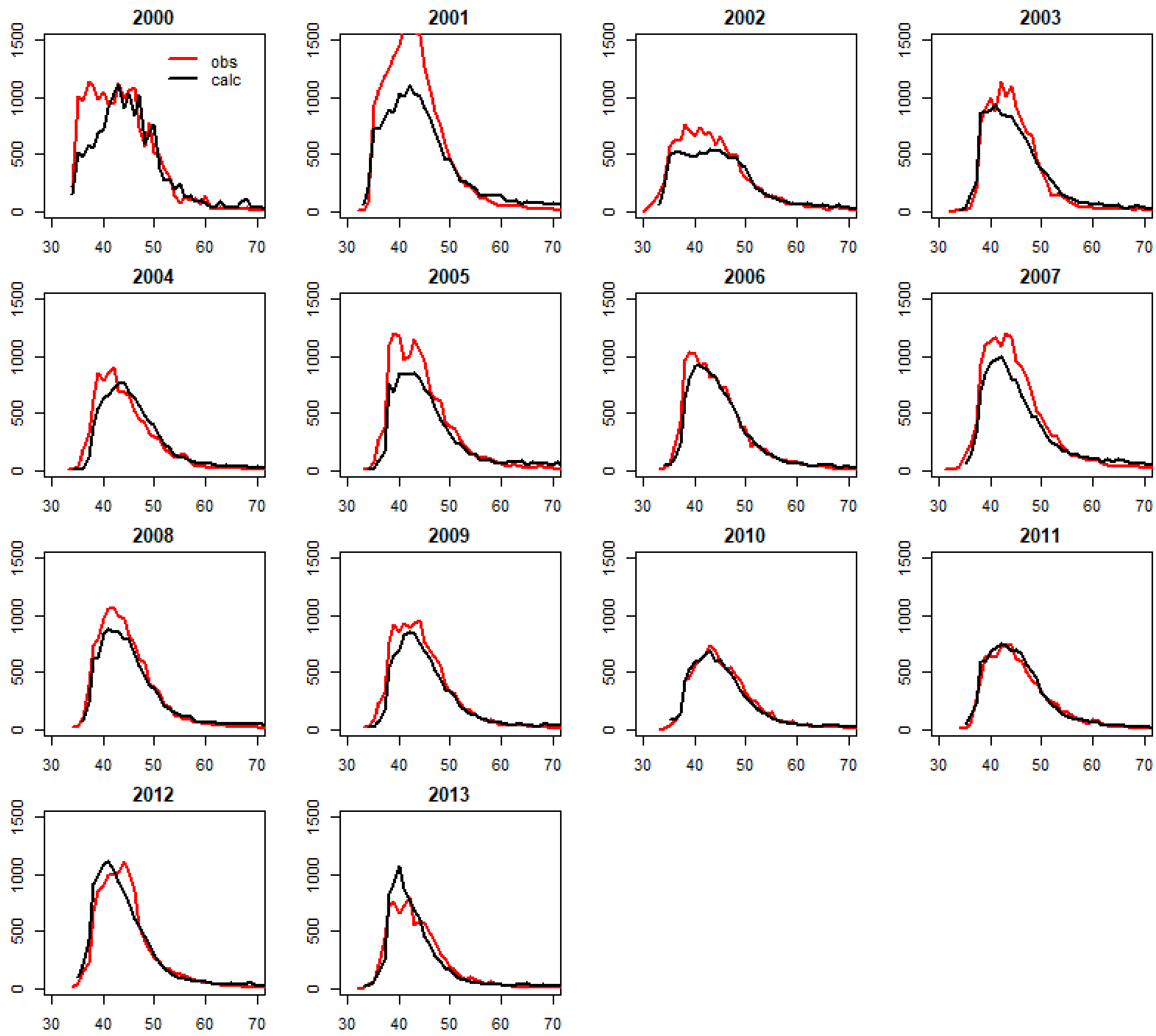


Fig 6. Observed and calculated landings at length in SD 24.

2 TIME SERIES OF WBC COMMERCIAL CATCH

Landings at age in SD 22-23

Landings at age for SD 22 were derived from the multispecies assessment databases for the Baltic Sea, and from 1996 onwards the landings at age by SD were available from WGBFAS reports. For SD 23, landings at age for 1997-2013 were derived from WGBFAS reports (Lindegren et al. 2013). For 1979-1996, the landings at age in SD 22 were upscaled by the landings taken in SD 23 compared to SD 22, to obtain landings at age for SD 22-23.

Discard in tons and number at age

The calculation procedures for different years are described in the table below.

Years	***A:*** SD 22-23 Tons	***B***: SD22-23 N@age	***C:*** SD24 Tons	***D:*** SD24 WBC Tons	***E:*** SD22-24 WBC N@age
2014-2017	From WGBFAS (IC)	From WGBFAS (IC)	From WGBFAS (IC)	***C*** multiplied by combined stock split key for SD 24	***B*** raised to account for tons in SD 24 (***D***)
2011-2013	Sumproduct of discards at age (***B***) and weight at age.	Annual discards at age in SD22-24 from former assessments adjusted with the average proportion of SD22-23 in the total discards in SD 22-24 in years 2008-2010.	Discard tons in SD22-24 from former assessments minus SD 22-23 (***A***)	***C*** multiplied by combined stock split key for SD 24	***B*** raised to account for tons in SD 24 (***D***)
2002-2010	Sumproduct of discards at age (***B***) and weight at age.	Discards at age in SD22 and SD23 from previous WGBFAS assessments, summed.	Discard tons in SD22-24 from former assessments minus SD 22-23 (***A***)	***C*** multiplied by combined stock split key for SD 24	***B*** raised to account for tons in SD 24 (***D***)
1996-2001	Sumproduct of discards at age (***B***) in SD 22 and weight at age.	Discards at age for SD22 from previous WGBFAS assessments.	Discard tons in SD22-24 from former assessments minus SD 22 (***A***)	***C*** multiplied by combined stock split key for SD 24	***B*** raised to account for tons in SD 24 (***D***)
1980-1995	Sumproduct of discards at age (***B***) in SD 22 and weight at age.	Extrapolated: landing at age in SD 22 multiplied by average discard ratio in 1999-2003.	Not included	Not included	***B***

BMS (below minimum reference size)

BMS is not estimated in the same way in all countries. In many countries the BMS fraction is included as part of the discard estimate derived from the observer trips. The reason for this is due to the handling process on board where it is not always obvious for the observer if a basket of fish is going to be discarded or if it is landed as BMS. However, in the ICES data call it is possible to upload the BMS data as official data (derived from the logbooks) and have a 0 as the estimated BMS fraction from the observer trips. If this is the case the official BMS number can be subtracted from the discard estimate to ensure the BMS is not double accounted. Denmark and Poland is presently not able to separate discard and BMS at the observer trips and BMS is therefor included in the discard estimate for these countries and the amount is afterwards subtracted with the official registered BMS. For Germany and Sweden BMS fraction is estimated from the observer trips.

Weight at age

Annual weight at age in landings in SD 22 was available from 1994 onwards, from earlier WGBFAS. For earlier years, average values for 1994-1998 were applied.

For SD 23, annual mean weight was available from 1997 onwards. For earlier years, average of 1997-1999 was applied.

For SD 22-23, data were available directly from IC for years 2014 onwards. For earlier years, mean of SD22 and SD23 was applied, weighted by landings at age. These are used to represent mean weight in landings of the entire WBC stock.

Weight at age in discards in SD22-23 were available for 2014 onwards. For earlier years, values equal to 2014 were applied. These are used to represent mean weights of total discards of WBC stock.

Weight at age in recreational catch…

Total weight at age in catch is derived by averaging the mean weights at age in landings, discards and recreational catch, weighted by respective catch numbers.

References

Hüssy, K., Mosegaard,H., Albertsen, C.M., Nielsen, E.E., Hemmer-Hansen,J., Eero, M. 2016a. Evaluation of otolith shape as a tool for stock discrimination in marine fishes using Baltic Sea cod as a case study. Fisheries Research, 174: 210-218.

Hüssy, K., Hinrichsen, H.-H., Eero, M., Mosegaard, H., Hemmer-Hansen, J., Lehmann, A., and Lundgaard, L. S. 2016b. Spatio-temporal trends in stock mixing of eastern and western Baltic cod in the Arkona Basin and the implications for recruitment. ICES Journal of Marine Science, 73 (2): 293-303 , doi: 10.1093/icesjms/fsv227.

WD 14 - German recreational catch estimation

Harry V. Strehlow & Simon Weltersbach, Thünen-Institute of Baltic Sea Fisheries

1. Description of the fishery

Recreational fishing is under the jurisdiction of the 16 German federal states. In case of the Baltic Sea these are Schleswig-Holstein (SH) and Mecklenburg-Western Pomerania (MV). Although recreational fishing licenses (does not distinguish between freshwater or saltwater fishing) are obligatory for fishing in the Baltic Sea only the state of MV demands an additional coastal fishing permit for the Baltic Sea allowing a direct estimation of number of anglers (Strehlow *et al.*, 2012). However, German data privacy protection legislation prevents the general use of personal data from license registries for research purposes (when purchasing a MV coastal fishing permit it is possible to indicate contact details for research purposes on a voluntary basis). Therefore, no representative sampling frames e.g. all German fishing license holders is available for recreational fisheries surveys (Strehlow *et al.*, 2012).

Recreational cod catches are mainly taken by private (60%) and charter boats (25%) and to a smaller degree by land-based fishing methods (15%) in recent years. Rod-and-line fishing with artificial lures or live bait is the primary fishing method targeting cod (Weltersbach *et al.*, 2019). Cod angling takes place throughout the year in German Baltic waters. Only few passive gear licenses exist and numbers are declining since many years. Furthermore, German recreational passive gear fishers mainly target eel and flatfish and cod catches are negligible (< 1% of the total recreational catches) and therefore not regularly estimated (Strehlow *et al.*, 2012). There are no seasonal or spatial closures regulating the marine recreational fishery for cod. The legal minimum landing size (MLS) for cod was 38 cm until 2016, since 2017 MLS is 38 cm in SH and 35 cm in MV. In 2017, a bag limit was introduced limiting recreational cod harvest to 3 cod per day/angler during the closed season (February to March) and 5 cod per day/angler for the rest of the year. In 2019 the bag limit was raised to 7 cod per day/angler for the entire year.

2. Description of the survey, data collection and catch estimation procedures

The German marine recreational fisheries data collection program follows a multiannual multistage survey design (for further information see Strehlow *et al.*, 2012) and has been established in 2004/2005. This multistage survey design involves the following components: (i) off-site survey (mail-diary / telephone-diary) for angling effort estimation, (ii) on-site survey (data from completed trips for a stratified random sample of access points and days) for catch-per-unit-effort (CPUE) estimation, (iii) recreational cod catch length distributions from length measurements, (iv) length-weight relationship and age-length keys from commercial and survey (BITS) cod catches for conversion of numbers into biomass and CATON and CANUM (Fig. 1).

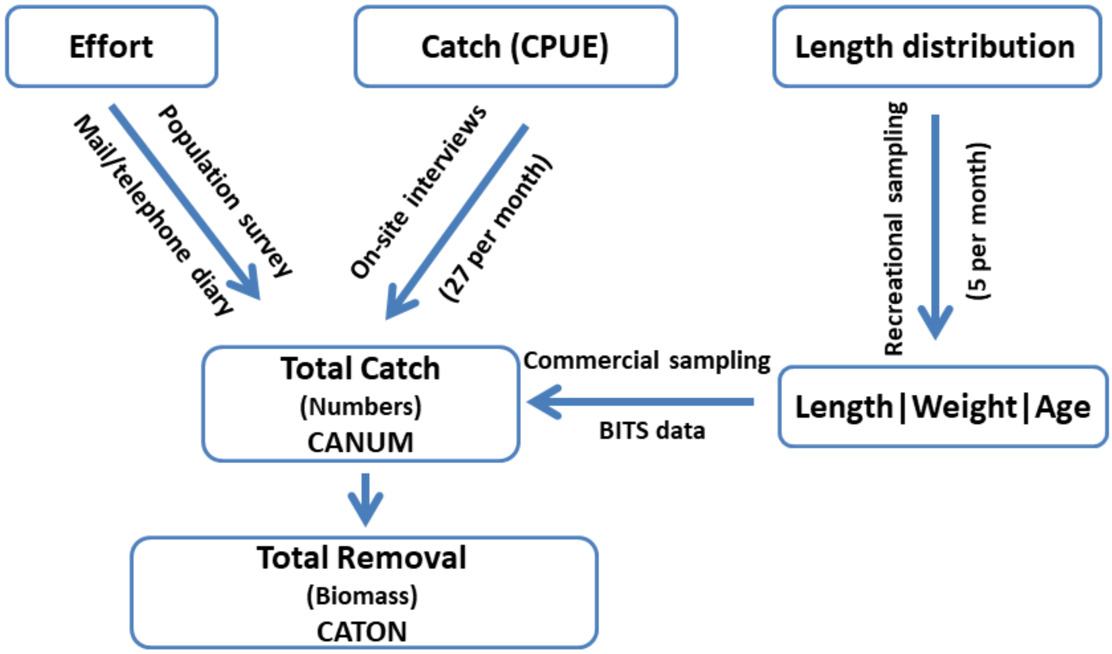


Fig. 1: Schematic overview of the German recreational cod catch survey and estimation procedure.

The marine recreational fisheries data are stratified into sea-based (boat angling (including belly boat and kayak angling), charter vessel angling and trolling) and land-based (shore angling and wading) fishing practices. CPUE data and recreational cod catch length measurements are collected on a monthly basis and grouped half yearly to reduce variance. Further stratification of the data is by coastal federal states (MV, SH), subdivision (22, 24), harvest and release component in numbers.

2.1 Number of anglers and angling effort

The annual angler population (annual number of German anglers fishing in the Baltic Sea) is estimated using the official coastal fishing permit sales data from MV. Anglers fishing without a valid license (illegal) are not accounted for but this problem is considered to be minor. Annual (2.9%) and weekly (3.7%) licenses are corrected to account for license holders fishing in MV and SH. Daily license sales numbers are corrected (daily licenses correction factor = 5) to account for the numbers of day licenses issued per angler per year. These correction factors are based on mail-diary respectively telephone-diary data. The combined corrected annual, weekly and daily license sales numbers equal the estimated number of anglers in MV. To estimate numbers of anglers in SH a conversion factor is used based on the proportion of state fishing license holders in MV and SH (Table 1). The corresponding conversion factor is then multiplied with the annual state fishing license sale numbers of SH assuming that the ratio of anglers that purchased a state fishing license in MV and fished in the Baltic Sea (coastal fishing permit holders) is the same in SH.

Table 1: Actual license sale numbers of coastal fishing permits in MV (annual, weekly and day licenses) from 2005-2017 and estimated numbers of German Baltic Sea anglers per coastal state and in total.

Year	Annual licenses MV	Weekly licenses MV	Daily licenses MV	# anglers MV	# anglers SH	Total anglers
2005	53,512	18,692	30,324	76,453	53,494	129,948
2006	53,703	22,679	34,599	81,327	61,988	143,315
2007	54,845	19,895	31,448	79,128	55,460	134,588
2008	56,022	20,623	27,347	80,175	56,776	136,951
2009	57,468	21,415	26,521	82,188	57,301	139,489
2010	57,341	19,192	18,708	78,360	48,504	126,864
2011	61,743	23,575	24,351	88,019	48,752	136,771
2012	61,607	24,722	25,833	89,287	60,447	149,735
2013	63,892	26,246	26,541	93,133	63,051	156,185
2014	57,682	33,710	33,686	95,671	64,769	160,440
2015	58,976	36,749	36,149	100,357	67,941	168,298
2016	57,994	39,084	36,774	101,769	68,897	170,666
2017	55,762	36,675	32,373	96,384	65,252	161,635

An off-site survey (mail-diary) from 2004/2005 was used to estimate angling effort (mean angling days per angler and year) until 2015. The main objective was to obtain effort data, i.e. how many days did an angler go fishing in the Baltic Sea and by which fishing method. This involved a mail-diary survey with 66,000 questionnaires of which 2,313 were evaluated to estimate numbers and fishing effort of anglers (see Strehlow et al. (2012) for details). Total annual angling effort was then estimated by multiplying the total number of Baltic Sea anglers per year (derived from the coastal fishing permit sales data from MV and the corresponding license sales estimates from SH; Table 1) with the mean angling effort per angler and year derived from the mail-diary survey (Table 2). Due to financial and logistical constraints (large German population (~ 82M) and an overall low participation rate (< 0.25 %) in sea angling) no update of the fishing effort estimates was possible until 2015 which means that the mean angling days per angler where held constant until 2015. However, due to consideration of the variance of the annual license sales data in the calculation procedure (as described above) some variation in the total angling effort can be accounted for.

In 2014/2015, a one-year national telephone-diary survey covering 9 out of 16 German federal states, with quarterly follow-ups was initiated to update the fishing effort data from the 2004/2005 mail-diary survey (Table 2). Thereby two federal states where used as proxies for the remaining 7 federal states not covered by the survey. This approach seemed reasonable as the 9 federal states covered by the survey accounted for ~75% of the total angler population fishing in the Baltic Sea and ~80% of the total angling effort in the Baltic Sea, respectively. Furthermore, data derived from on-site interviews revealed that the selected two federal states were a good representation of the 7 federal states not covered by the survey in terms of participation rates and angling characteristics. In total, 358,411 telephone numbers were generated by random digit dialing, which resulted in a net sample of 50,200 valid telephone numbers of private households. 678 marine anglers were identified and 348 panellists recruited. The panel was later boosted using a non-representative sample of MV coastal fishing permit holders resulting in 582 additional panellists (Hyder et al. 2018). Since 2015 effort estimates (mean angling days per angler and year) are based on this nationwide telephone-diary survey (Table 2).

Table 2: Mean angling effort (d) per angler and year divided by federal state, half year, fishing method and license type (annual, weekly, daily) for the two population surveys conducted in 2004/2005 (mail-diary) and 2014/2015 (telephone-diary).

				Mean angling effort (d) per angler and year		
Survey year	Federal state	Half year	Method	Annual license holders	Weekly license holders	Daily license holders
2004/2005	MV	first	Boat	1.10	0.17	0.78
2004/2005	MV	second	Boat	1.36	0.44	1.03
2004/2005	MV	first	Charter vessel	0.81	0.17	0.28
2004/2005	MV	second	Charter vessel	0.91	0.28	0.45
2004/2005	MV	first	Shore	1.26	1.17	0.16
2004/2005	MV	second	Shore	1.31	0.39	0.38
2004/2005	SH	first	Boat	1.24	NA	NA
2004/2005	SH	second	Boat	2.08	NA	NA
2004/2005	SH	first	Charter vessel	0.62	NA	NA
2004/2005	SH	second	Charter vessel	0.63	NA	NA
2004/2005	SH	first	Shore	2.84	NA	NA
2004/2005	SH	second	Shore	3.41	NA	NA
2014/2015	MV	first	Boat	2.29	1.06	0.37
2014/2015	MV	second	Boat	1.35	0.34	1.14

2014/2015	MV	first	Charter vessel	0.22	0.40	0.29
2014/2015	MV	second	Charter vessel	0.46	0.31	1.20
2014/2015	MV	first	Shore	1.37	0.61	0.86
2014/2015	MV	second	Shore	1.34	1.89	0.63
2014/2015	SH	first	Boat	1.42	NA	NA
2014/2015	SH	second	Boat	1.59	NA	NA
2014/2015	SH	first	Charter vessel	0.35	NA	NA
2014/2015	SH	second	Charter vessel	0.53	NA	NA
2014/2015	SH	first	Shore	1.28	NA	NA
2014/2015	SH	second	Shore	1.80	NA	NA

Changes between the mail-diary and the telephone-diary survey are mainly between platforms, i.e. a decrease in charter boat effort and increasing effort from boats (Table 2). However, changes in average fishing effort are marginal.

2.2 Catch-per-unit-effort data

On-site, a stratified random sample of access points – (87 access points (beaches, harbours, and piers) covering the German Baltic coast from Flensburg to the island of Rügen) – and days is used to interview anglers (only completed trips) and estimate catch rates (CPUE) per stratum (Figure 2).

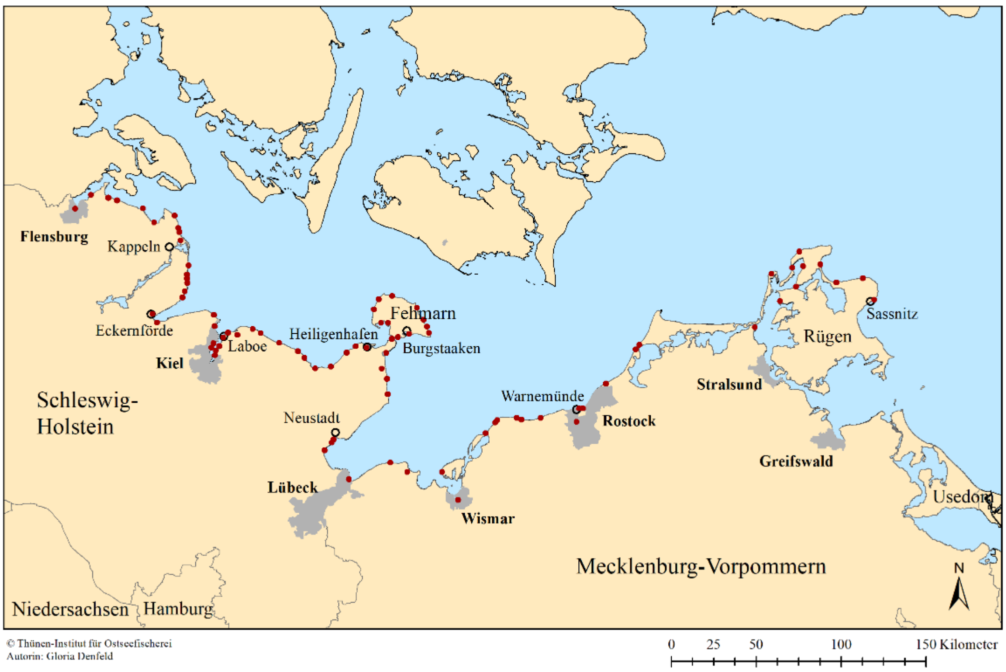


Figure 2. Map of access point (beaches, harbours, and piers) for the German multiannual stratified random intercept survey of the Baltic Sea to sample boat-based (boat, charter vessel) and land-based (wading, shore fishing) fishing activities.

On-site sampling is carried out as peak activity sampling without replacements during 27 days per month by six survey agents (Table 3). Therefore, access points are sampled during the period where the probability of returning anglers from sea is highest. Observation time per access point is usually 2-3 hours. Land-based fishing is sampled during the typical fishing time in the evening hours.

Table 3: Overview of the numbers of on-site surveys and interviewed anglers, 2005 – 2017.

Year	Number of on-site surveys	Numbers of interviews
2005	183	1494
2006	168	2386
2007	162	1888
2008	129	1052
2009	253	2288
2010	290	2359
2011	341	2816
2012	316	1983
2013	324	1587
2014	315	1535
2015	323	1469
2016	329	1604
2017	324	1392

2.3 Post-release mortality

A large amount of cod is released alive voluntary or regulatory due to minimum landings sizes and bag limits (Ferter et al. 2013). Anglers release fish with the assumption that they will survive. Several studies exist that have estimated post-release mortality of released cod (Capizzano et al. 2016; Ferter et al. 2015; Weltersbach & Strehlow, 2013). The amounts of dead recreational releases (discards) are estimated following two compilations methods:

- Dead land-based releases are estimated assuming 100% post-release mortality rate of all cod released in land-based fisheries (precautionary approach as no studies are available).
- Dead sea-based releases are estimated applying a 11.2% post-release mortality rate to all sea-based releases based on a post-release mortality study from 2012 (Weltersbach & Strehlow, 2013).

2.4 Length, weight and age data

Length distributions of recreational catches (harvest & release component) are collected during onboard measurements by survey agents on charter vessel trips (random selection of vessels and dates) five times per month stratified along the German Baltic coast (Table 4). Other data sources are self-reported length samples from fishing events e.g. for shore fishing. Commercial/BITS length-weight relationships and age-length keys are used for conversion of recreational catch numbers to biomass and length-at-age. This information is than used to derive CANUM and CATON for the assessment.

Table 4. Overview of the number of samples and length measurements of cod from recreational fishing events (charter vessels trips & shore fishing), boat and trolling self-measurements, as well as charter vessel sampling, 2005 – 2017 (all data pooled).

Year	Samples	# cod measured (harvest)	# cod measured (release)
2005	17	1461	
2006	6	362	
2007	6	516	8
2008	32	620	33
2009	87	1354	895
2010	87	3634	1635
2011	80	4673	1102
2012	32	1546	533
2013	47	2257	1345
2014	45	3721	1104
2015	42	2853	949
2016	53	2521	398
2017	45	937	1269

2.5 Reconstruction of recreational cod catches before 2005

German recreational cod catch data are available from 2005 continuously on an annual basis. To meet the demands of the assessment a reconstruction of the recreational cod catches from 1980 to 2004 was necessary (Table 5). Therefore, the average catch from 2009-2015 was used to extrapolate the years 1980 to 2004. To account for the historic development of marine recreational fishing in the former GDR after reunification, recreational catches in MV were set at 20% from 1980 - 1991 with a linear increase by 20% each year between 1991 to 1995. To convert recreational catches in numbers to CANUM the pooled recreational length distribution from 2005-2017 is used for the years 1980 to 2004. From 2005 onwards CANUM is estimated based on the respective recreational length distribution from each year. The ALK used for conversion are based on BITS data for the years 1980 to 2002 and on commercial sampling data from 2002 onwards. When stratified length distributions are missing, data is borrowed from adjacent strata.

Table 5: Overview of the data used for constructing the German recreational cod catch time series from 1980-2017 for the assessment.

Time series	1980-2002	2002-2004	2005-2014	2015-present
# angler	Ø 2005-2015 from license sales in MV corrected with mail-diary study		Annual license sales from MV corrected with mail-diary study	Annual license sales from MV corrected with mail/telephone -diary
# days	Ø 2005-2015 from mail-diary study		mail-diary study	telephone-diary study
CPUE	na	na	on-site intercept surveys	
WLK & ALK	BITS data	Commercial sampling		

2.6 CANUM & CATON

To compile catch-at-age the estimated recreational cod removals (harvest & dead releases) in numbers are distributed according to the recreational length distribution and proportions-at-length in the sample. Stratification is by SD, half-year, fishing mode (sea-based, land-based,) harvest and retained. The total numbers-at-length (recreational) are then matched with the ALK from German commercial fisheries or BITS data for each year. Age proportion per length is used to group numbers-at-length into numbers-at-age categories stratified by SD and half year. Based on a decision during WKBALTCOD 2015 all recreational catches are considered WBC, accordingly commercial and BITS data from SD22 is used to produce CANUM and CATON from recreational catches in SD24. In case of insufficient length categories, we use an alternative basic sample of pooled commercial data (> 2009).

Weight estimation follows a similar approach as explained above except that the length mass coefficient of the LW relationship is used to calculate a mean-weight-per-length-class. This mean-weight-per-length-class is summed up and divided by numbers-at-age to calculate mean-weight-at-age.

3. Changes since last benchmark

Several changes have been made since WKBALTCOD 2015 that affect the German recreational cod catch time series. These changes comprise:

- Minor adjustments of the input data due to systematic search of errors in the database
- New effort estimations (mean angling days per angler and year; Table 2) based on the telephone-diary survey conducted in 2014/2015 from 2015 onwards.
- Extension of the time series from 1991 to 1980 to meet assessment needs.

The total change in the CATON time series varies between -29% and +37% (Figure 3). On average CATON increased by +7%.

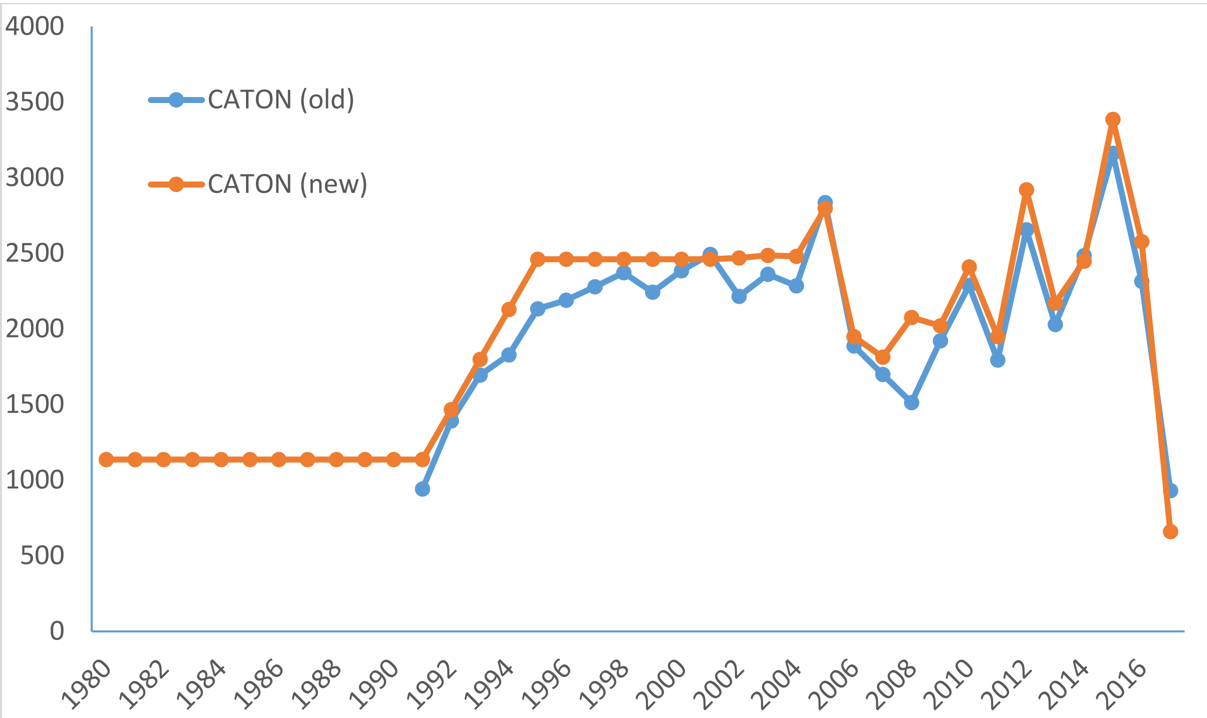


Figure 3. Changes in German recreational CATON time series of WBC in tons old versus new data.

4. References

Capizzano, C. W., Mandelman, J. W., Hoffman, W. S., Dean, M. J., Zemeckis, D. R., Benoît, H. P., Kneebone, J., et al. 2016. Estimating and mitigating the discard mortality of Atlantic cod (Gadus morhua) in the Gulf of Maine recreational rod-and-reel fishery. ICES Journal of Marine Science: Journal du Conseil, 73: 2342-2355.

Ferter, K., Weltersbach, M. S., Humborstad, O. B., Fjelldal, P. G., Sambraus, F., Strehlow, H. V., and Volstad, J. H. 2015. Dive to survive: effects of capture depth on barotrauma and post-release survival of Atlantic cod (*Gadus morhua*) in recreational fisheries. ICES Journal of Marine Science, 72: 2467-2481.

Ferter, K., Weltersbach, M. S., Strehlow, H. V., Vølstad, J. H., Alós, J., Arlinghaus, R., Armstrong, M., et al. 2013. Unexpectedly high catch-and-release rates in European marine recreational fisheries: implications for science and management. ICES Journal of Marine Science: Journal du Conseil, 70: 1319-1329.

Hyder, K., Weltersbach, M. S., Armstrong, M., Ferter, K., Townhill, B., Ahvonen, A., Arlinghaus, R., et al. 2018. Recreational sea fishing in Europe in a global context—Participation rates, fishing effort, expenditure, and implications for monitoring and assessment. Fish and Fisheries, 19: 225–243.

Strehlow, H. V., Schultz, N., Zimmermann, C., and Hammer, C. 2012. Cod catches taken by the German recreational fishery in the western Baltic Sea, 2005 - ì2010: implications for stock assessment and management. ICES Journal of Marine Science: Journal du Conseil, 69: 1769–1780.

Weltersbach, M. S., and Strehlow, H. V. 2013. Dead or alive—estimating post-release mortality of Atlantic cod in the recreational fishery. ICES Journal of Marine Science: Journal du Conseil, 70: 864–872.

Weltersbach, M.S., Lewin, W.-C., Gröger, J.P., Strehlow, H.V. 2019. Effect of lure and bait type on catch, size, hooking location, injury and bycatch in the western Baltic Sea recreational cod fishery. Fisheries Research, 210: 121–130.

WD 15 - Documentation of data collection and analysis

Table of Contents

Documentation of data collection and analysis

2019-01-04

Introduction

This document describes the methods used to calculate the Danish recreational harvest of Western Baltic cod in ICES SD 22, 23 and 24. The estimates are based on data gathered from two different ongoing surveys covering the recreational fisheries; an off-site survey (DST) and an on-site survey (REKREA) and an omnibus survey conducted in 2010 and 2011. The off-site survey is a biannual interview-based recall survey which has been running since 2009 while the on-site survey was introduced in 2016. Knowing that respondents may overestimate the effort in recall surveys (Sparrevohn and Storr-Paulsen 2012) the discrepancies in the calculated harvest in the two surveys were used to adjust the calculated harvest for the DST survey going back to 2009. The adjusted harvest was used as a basis for hindcasting the estimates for cod harvest back to 1980 for ICES SD 22, 23 and 24.

DST survey

DST is a biannual recall based web survey carried out by Statistics Denmark together with DTU Aqua. It targets recreational fishers holding a valid 1-year license at the time of the interview. Since two different license lists are available, one for anglers and one for passive gear fishers, two surveys are conducted with similar questionnaires. Independent of list, the respondents are randomly selected (simple random sample, SRS) and initially 2500 – 3500 fishers of each license type are contacted by letter in each biannual survey wherein they are encouraged to answer the questions via the internet. Respondent rate is typically between 30 – 45% (Table 1). The questionnaire contains detailed questions on harvest and release of cod and fishing effort within the last 6 months. The respondent is explicitly told to distinguish between the part of the catch kept (i.e. the harvest) and the part released (discarded). To estimate harvest by ICES managing areas (Fig. 1) and quarter the respondents are asked to provide the information per area and quarter. Since 2016, respondents have been asked about the number of trips in the Sound (SD23) on private boats, tour boats or by the coast, and how much cod (kg) they have caught from these different platforms.

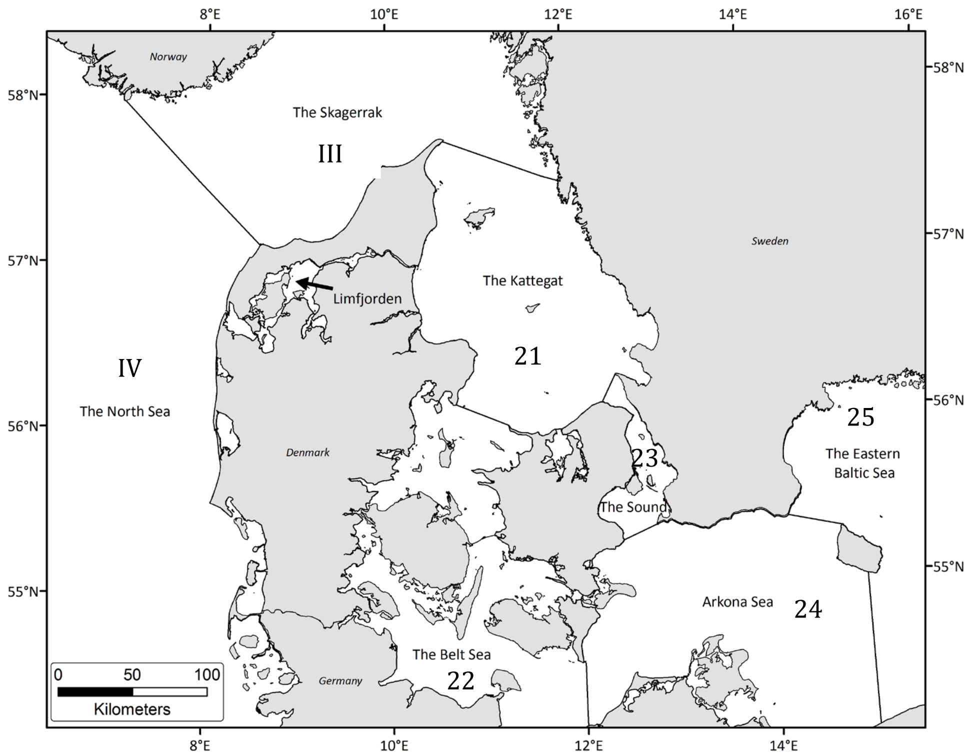


Figure 1. Map showing the ICES areas for which respondents are asked to report their catches of cod.

Respondents can report their harvest as weight or numbers. If harvest is reported in numbers they have to be transformed to weight estimates multiplied with an average fish weight (See under 'Number to weight').

In the Danish license system it is also possible to issue a license valid for one day or one week. However, the number issued of these licenses is relatively small compared to the number of annual licenses. Therefore, no separate interviews are conducted for these two groups. However, they are accounted for in the total harvest estimations, taking the different effort into account. Furthermore, the purchasing a license for passive gear fishing automatically gives license to angle with rod and reel as well. To include this group in the estimates, all passive gear fishers are asked whether he/she also angled. An additional interview is therefore conducted on this group in order to estimate their harvest when angling.

Table 1. DST survey participation for 2016

DST survey (questionnaire) 2016	Anglers		Passive gear fishers	
	1. halfyear	2. halfyear	1. halfyear	2. halfyear
Number of fishers in subsample	2810	3133	2524	2926
Number of respondents	1166	875	1100	1239
Respondent rate (%)	41.5	27.9	43.6	42.3

Omnibus survey

The main objective of this interview was to estimate the size of the population that fished without a license and with what effort. The Omnibus is a monthly survey conducted by Statistic Denmark wherein questions are asked on behalf of e.g. companies, newspapers and research institutes. In 2009, three telephone interview rounds were conducted were questions on recreational fishery were included and in 2010 one additional omnibus survey was conducted in March. The recreational fishery questions were embedded as a minor part of this interview; hence the non-response bias is expected to be ignorable. Respondents were selected by telephoning a random number. The interview was conducted with that person within the household who last had birthday. Only citizens between 16 and 74 were included. A total of 958, 957 and 968 were interviewed and answered in 2009 and in March 2010 a total of 985 were interviewed.

Analytical methods DST

Estimating the total harvest or numbers released of cod was done by estimating the harvest on basis of the reported catches from the license list recall survey. These values were then extrapolated to the entire population of fishers (all license holders and fishers without a license) using effort information collected during the omnibus survey. Different effort levels for those fishing without a license, on a weekly or on a daily license were accounted for in the calculation. To compute the total harvest or released numbers $\hat{Y}_{ij}$ of cod per quarter (i) and area (j) the following equation was used:

$$\hat{Y}_{ij} = \frac{\sum_{k=1}^{n_{ij}} y_{ijk}}{n} N$$

where n is the number of respondents and y the reported harvest per respondent (k). The total population N is computed as:

$$N = \left(\rho_a + \rho_w \cdot \frac{\varepsilon_w}{\varepsilon_a} + \rho_d \cdot \frac{\varepsilon_d}{\varepsilon_a} + \rho_m \cdot \frac{\varepsilon_m}{\varepsilon_a}\right)$$

where ρ is the number licenses issued being valid for a year (a), week (w) or day (d). The number fishing without a license (m) was computed using the estimated percentage that fished without a license even though obliged to have one, multiplied with the actual number of Danish citizens between age 18 and 65. The values were weighted with the fishing effort (ε) which for those holding an annual license was derived from the omnibus survey and assumed to be 1 day for those holding a daily license and 3 days for those holding a weekly license.

Number to weight

If the respondents report a number of cod, the number is multiplied by the value codMultHome, which is:

$codMultHome = 1.5$

This is the number to kg multiplier, which is used in earlier studies on the cod in The Sound (Sparrevohn and Storr-Paulsen 2010).

If the respondents report a number of released cod in the Sound, this number is multiplied by the value codMultRel, which is:

$codMultRel = 0.32$

This value is the median weight of cod discarded in the Sound. The data come primarily from commercial vessels and BITS surveys in SD23.

These values will be used for converting numbers to amounts of cod in both the DST and REKREA surveys.

Comparing weights

Ideally, respondents would report the same weight of caught cod in the Sound in general and as a total on the three platforms. This is not always the case. To compensate for this discrepancy, we find a correction value based on the mean relation between the reported weight on the Sound and the sum of weights on the 3 platforms for each respondent (Fig.2). We call this correction value *harvestCorrection*.

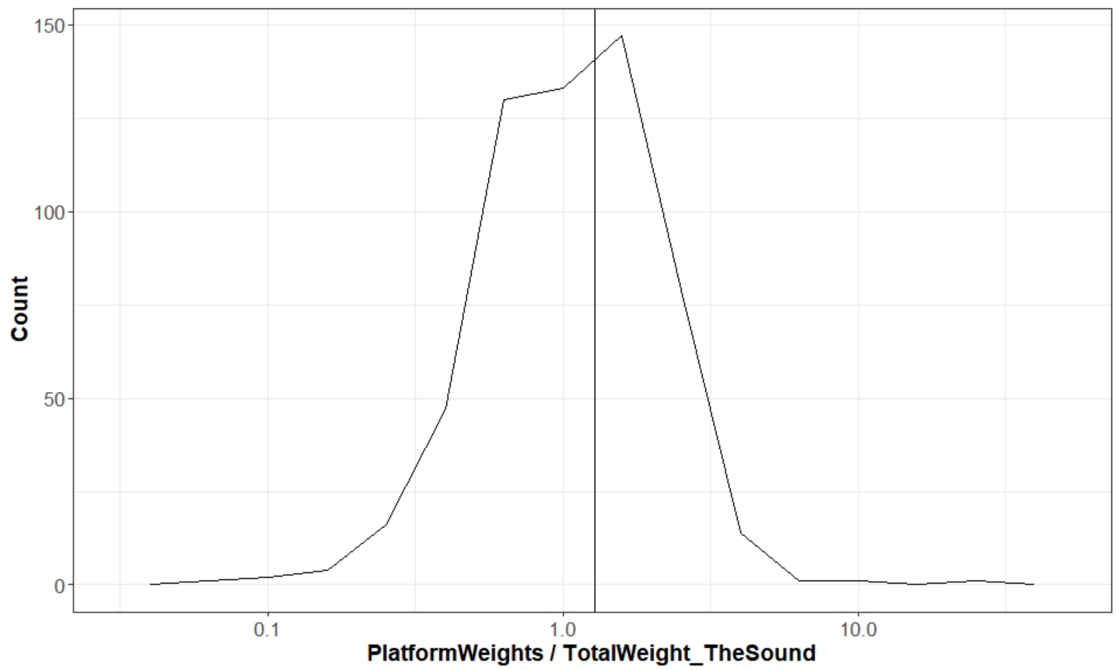


Figure 2. Relation between reported weights on the three platforms and the total weight reported for the Sound for each respondent

$$harvestCorrection = \frac{\sum_{i=1}^{n} \frac{w_{c,i} + w_{tb,i} + w_{pb,i}}{w_{x23,i}}}{n}$$

where $w_{c,i}$ is the reported weight of cod caught from the coast, $w_{tb,i}$ is the reported weight of cod from tour boats, $w_{pb,i}$ is the weight of cod caught from private boats, $w_{x23,i}$ is the total weight reported in the Sound before specifying the weight per platform, and n is the number of respondents.

HarvestCorrection will be used to correct the catches from the different platforms in order to be comparable with the reported weights from the Sound in the DST survey back to 2009:

$$w_{corrected} = \frac{w_{coast} + w_{tourboat} + w_{privateboat}}{harvestCorrection}$$

The corrected weights and number of trips reported by respondents are grouped per year, platform and quarter.

Effortmultiplier

The weights from the survey are scaled up using an effort multiplier and converted to tonnes. The effortMultiplier is found for each of the three platforms and in all quarters.

$$effortMultiplier_{p,q} = \frac{Trips_{p,q} * Licenses_y}{sampleSize_h * 1000}$$

where $Trips_{p,q}$ is the number of trips on a platform in a quarter, $sampleSize_h$ is the number of respondents in the DST survey of the half year that covers the quarter and $Licenses_y$ is the total number of fishing licenses in the given year.

Harvest per quarter

The harvest in the DST survey can now be found with the following formula:

$$Harvest_{p,q} = \frac{\sum_{i=1}^{n} \frac{weight_{i,p,q}}{trips_{i,p,q}}}{1000} * effortMultiplier_{p,q}$$

Where n is the number of respondents in the survey, and p is the platform from which the cod has been caught and q is the quarter.

Rekrea survey

Rekrea is an on site interview survey performed on tour boats and harbours along the Danish coast of the Sound. Observers from DTU Aqua interview anglers onboard tour boats or when private boats return to harbours. The survey is primarily providing catch per angler per trip and biological data for cod i.e. length, weight and age. The interviews are carried out using tablets. Answers are uploaded by the observers to an online survey platform when they return from the harbour. The data is later downloaded for analysis. In the analysis, tables are generated with trip information and catch information.

Design

Table 2. Sampling design for REKREA

Variable	**Description**
psu	Fishing trip
psu_strata (h)	Quarter
psu_prob (p)	Selection probability

n	Sampled number of fishing trips per platform per strata
ssu	Angler
M	Total number of anglers within PSU
m	Sampled number of anglers within PSU
harvest	UPSWOR unequal probability sampling without replacement
PSU_selection_method	The estimators are much more simple if with replacement can be assumed. We sample <10% of the fishing trips
SSU_selection_method	SRS simple random sampling. Mostly census.

Sampling effort

Table 3 showing the sampling and number of respondents interviewed.

		Charter boat	**Charter boat**	**Private boat**	**Private boat**
Year	**Quarter**	**Number of boats sampled**	**Sum of respondents**	**Number of harbours sampled**	**Sum of respondents**
2016	3	6	45	1	4
2016	4	7	77		
2017	1	3	38	1	0
2017	2	12	91	7	6
2017	3	13	139	1	16
2017	4	9	60	3	14
2018	1	10	131	13	10
2018	2	13	169	25	42
2018	3	22	201	45	34
2018	4	13	154	38	15

Selection of PSU

The selection of PSU (tour boats or harbours) is probability proportion to size based (PPS) (see Table 2). The selection of tour boats is based on the reported number of anglers per quarter pr. tour boat. The selection of harbours is based on the number of private boat trips i.e. boats launched.

Anglers that wants to participate in the survey are interviewed regardless of their target species on the trip and regardless if they have caught anything or not.

Catch information

Anglers are asked about their catch and releases on the specific trip. With the anglers's permission, the available cod are measured and weighted and otoliths are sampled for age reading. In cases where only the length is known, the weight can be calculated by creating a linear model on the relation between $log_{10}(length)$ and $log_{10}(weight)$ for registered cod with known lengths and weights

$$weight_{est.} = 10^{(a*log_{10}(l)+b)}$$

where l is the known length of a cod, a is the slope in the linear model, in this dataset 2.963114, and b is the intercept value of the linear model, in on this dataset -4.97074.

Anglers are also asked about how many fish they have released on the trip. This number is converted to kilogram using the *codMultRel* value that is also used in the DST survey.

Estimating harvest per angler per strata

We find the mean harvest per angler per platform and quarter by looking at the relation between the weight, y_h and anglers x_h in each strata.

$$\overline{y}_h = \frac{1}{n_h}\sum_{i=1}^{n}\frac{y_{i,h}}{p_{i.h}}$$

where n_h equal to number of sampled fishing trips on a platform in a quarter, $y_{i,h}$ is the total weight on each trip and $p_{i,h}$ is the propability of the tour boat or harbour to be selected in that quarter.

$$\overline{x}_h = \frac{1}{n_h}\sum_{i=1}^{n}\frac{x_{i,h}}{p_{i,h}}$$

where $x_{i,h}$ is the number of respondents on the trip.

With these values, we can calculate the rate $\hat{B}_h$ and it´s associated variance within strata

$$\hat{B}_h = mean\ harvest\ per\ angler_h = \frac{\overline{y}_h}{\overline{x}_h}$$

$$v\hat{a}r\ (\hat{B}_h) = \frac{1}{(n_h)(\overline{x}_h)^2} \sum_{i=1}^{n} \frac{\left(\frac{y_i}{p_i} - \hat{B}_h \frac{x_i}{p_i}\right)^2}{n_h - 1}$$

Combining rate with inverse variance weighting

The weighting (w_h) of the harvest per angler (x_i) is related to the variance of the mean:

$$w_h = \frac{\frac{1}{v\hat{a}r\left(\hat{B}_h\right)}}{\sum_{h=1}^{n} \frac{1}{v\hat{a}r\left(\hat{B}_h\right)}}$$

where h is the quarter.

$$Weighted\ \hat{B} = \sum_{h=1}^{h_{max}} \hat{B}_h w_h$$

$$Weighted\ v\hat{a}r\left(\hat{B}\right) = \sum_{q=1}^{q_{max}} \sum_{p=1}^{p_{max}=3} v\hat{a}r\left(\hat{B}_h\right) w_h$$

Where p is the platform and q is the quarter.

Finding harvest per strata

To find a harvest value in each strata, we muliple the mean harvest per angler with the number of trips and the effortMultiplier, which we found in the DST survey data.

$$Harvest_{p,q,REK} = \frac{Weighted\ \hat{B}_{p,q,REK} * Trips_{p,q,REK}}{resp_{p,q,REK}} * effortMultiplier$$

where $Trips_{p,q,REK}$ is the number of trips (and respondents) on platform p in quarter q in the Rekrea survey, $resp_{p,q,REK}$ is he number of respondents on platform p in quarter q in the Rekrea survey.

$Weighted\ \hat{var}\left(\hat{B}\right)$ is be multiplied the same way:

$$Var_{Harvest_{p,q}} = Weighted\ \hat{var}\left(\hat{B}\right) * Trips_{p,q,R} * effortMultiplier$$

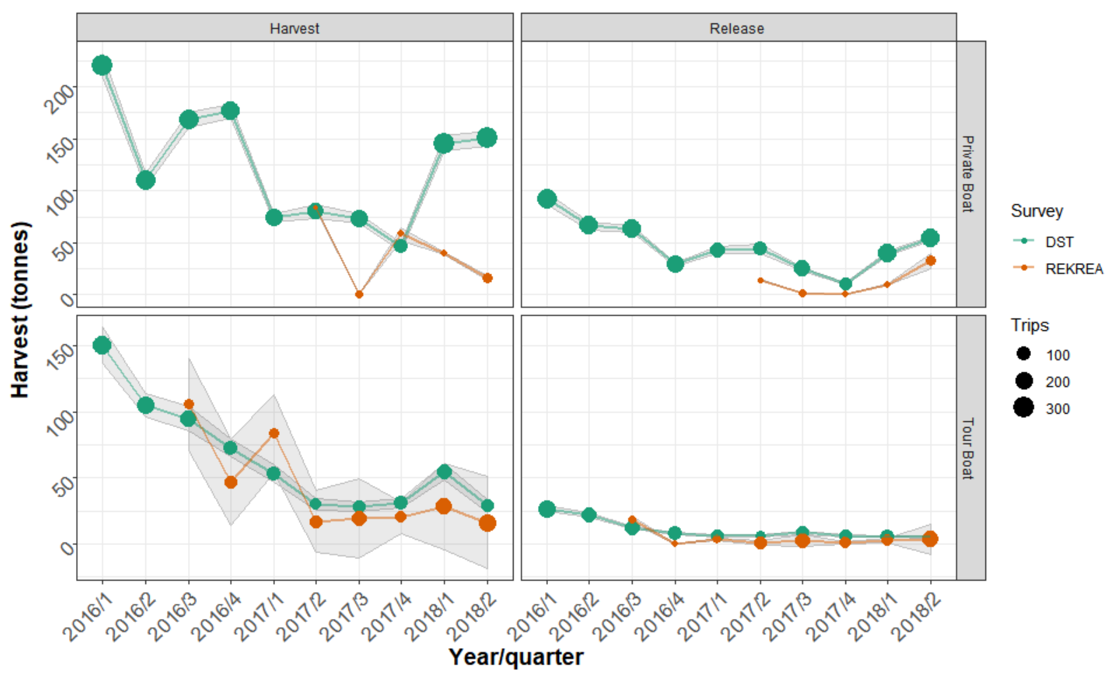
The harvest can be seen in the plot below with the variance in grey.

A standard error can be found:

$$SE_{p,q,s} = Licences_y * \frac{\sqrt{\frac{(totalResp_h - 1)^{-1} * (weight2_{p,q,s} - (weight^2_{p,q,s}/totalResp_h)))}{(totalResp_h) * (1 - totalResp_h/Licences_y)))}}}{1000}$$

where p is platform, q is quarter, s is survey, $totalResp_h$ is the total number of respondents in the DST survey in a given half year h, $weight2_{p,q,s}$ is the sum of each weight squared, $weight^2_{p,q,s}$ is the sum of all weights squared and $Licences_y$ is the number of licenses in a year.

Using the harvest numbers for each survey and platform, we can create the following graph.



Harvest per quarter per platform for DST and REKREA

The graph shows the calculated harvest and release per platform per quarter per survey.The numbers can be seen in the table below.

Year/quarter	Harvest/Release	Private Boat DST	Private Boat REKREA	Tour Boat DST	Tour Boat REKREA
2016/3	Harvest	168.2		94.7	105.7
2016/4	Harvest	176.8		72.6	46.8
2017/1	Harvest	74.3		53.0	83.4
2017/2	Harvest	80.0	83.1	29.9	16.8
2017/3	Harvest	73.4	0.0	28.1	19.4
2017/4	Harvest	46.1	58.6	30.7	20.0
2018/1	Harvest	145.5	39.8	54.6	27.9
2018/2	Harvest	150.2	16.2	28.7	16.2
2016/3	Release	63.5		12.1	18.4
2016/4	Release	29.1		7.9	0.3
2017/1	Release	43.0		5.8	3.4
2017/2	Release	44.1	14.0	6.2	1.0
2017/3	Release	25.0	1.4	8.4	2.6
2017/4	Release	10.9	0.0	6.1	0.8
2018/1	Release	39.4	9.7	5.4	2.9
2018/2	Release	54.1	32.5	5.1	3.7

RekreaMultiplier

The difference between the weights reported by respondents in the DST survey and the observed weights in the REKREA survey can be used to find another multiplier, called rekreaMult. Weights per area in the DST survey should be multiplied with in order to better fit what has been observed.

$$rekreaMult = \frac{\sum_{q=1}^{q_{max}} \left(\frac{H_{DST,q}}{H_{Rekrea,q}}\right)}{q_{max}} * harvestCorrection$$

Where q is the quarters with harvest data from both DST and REKREA, H is the harvest in tonnes calculated from the surveys DST and REKREA in the given quarter. As of January 2019, $q_{max} = 8$ expected to change to 10 in the spring of 2019.

The rekreaMultiplier is 0.6747 for harvest and 0.3463 for released cod. These numbers are used to as multipliers on the harvest and release numbers of cod in ICES areas 22, 23 and 24.

Adjusting old data

Going back to 2009, the harvest in the Sound, Western Baltic and the Belt Sea can be multiplied with the rekreaMult. Before 2009 we don't have data about the recreational fishing in the areas. Instead, we find a mean harvest per 18-65-year-old Dane (Statistik 2018), based on the harvest and release data from the years 2009-2017.

$$harvestPerAdult_{a,hr} = \frac{\left(harvest_{a,hr,y} * rekreaMult_{hr}\right)}{\sum_{p=18}^{p=65} P_{p,y}}$$

where P is the Danish population of people with a certain age each year, a is the area (22, 23 or 24) and hr is the harvest or release and y is the year.

$$meanHarvestPerAdult_{a,hr} = \frac{\sum_{y_{min}}^{y_{max}} harvestPerAdult_{a,hr}}{(y_{max} - y_{min})}$$

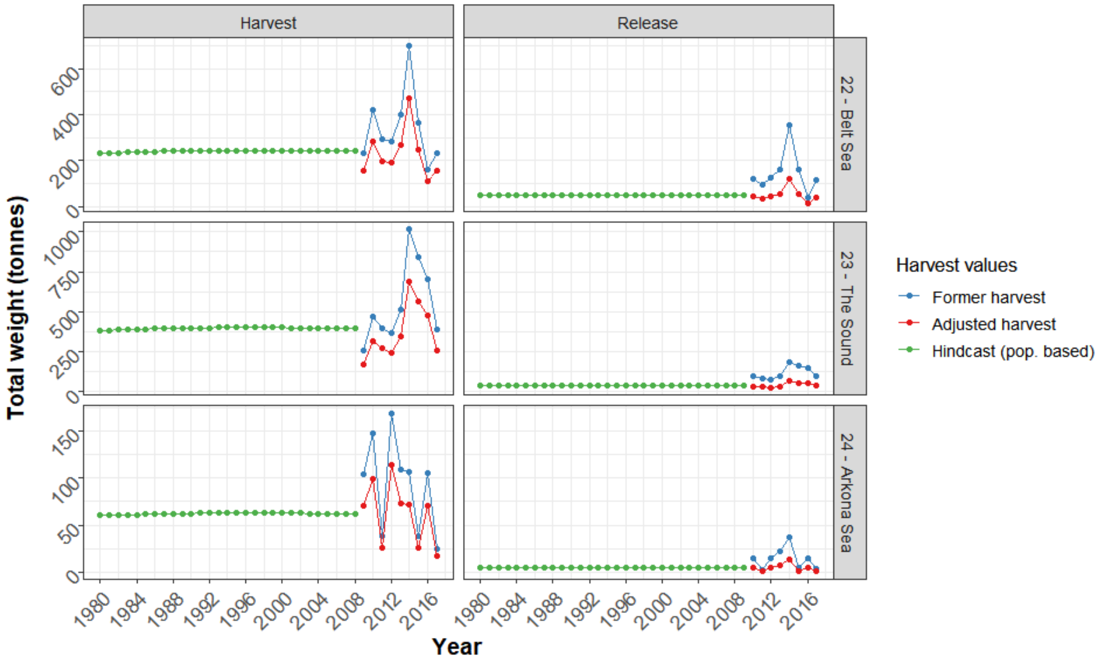
Where y_{min} is 2009, wher the DST survey began, y_{max} is 2017, the most recent full year we have survey data from.

The table below shows the harvestPerAdult values, we multiply the Danish population of 18-65-year-olds each year for each of the three areas.

Harvest/Release	Area	Mean harvest per adult
Harvest	22 - Belt Sea	7.1866e-05
Harvest	23 - The Sound	1.1792e-04
Harvest	24 - Arkona Sea	1.8537e-05
Release	22 - Belt Sea	1.5224e-05
Release	23 - The Sound	1.2194e-05
Release	24 - Arkona Sea	1.4640e-06

Adjusted harvest per year

The plot below shows the adjusted harvest in tonnes for harvest and release in Ices areas 22, 23 and 24.



Total harvest in the ices areas 22-24 before and after rekreaMult

Year	Source	22 - Belt Sea	23 - The Sound	24 - Arkona Sea
1980	Hindcast (Pop. based)	238.7995	387.1902	60.70040
1981	Hindcast (Pop. based)	239.4341	388.2192	60.86171
1982	Hindcast (Pop. based)	240.4503	389.8669	61.12003
1983	Hindcast (Pop. based)	241.6541	391.8187	61.42602
1984	Hindcast (Pop. based)	243.0335	394.0552	61.77664
1985	Hindcast (Pop. based)	244.6560	396.6860	62.18907
1986	Hindcast (Pop. based)	245.7739	398.4986	62.47324
1987	Hindcast (Pop. based)	246.5380	399.7375	62.66746
1988	Hindcast (Pop. based)	246.9845	400.4614	62.78094
1989	Hindcast (Pop. based)	247.2185	400.8408	62.84043
1990	Hindcast (Pop. based)	247.7559	401.7122	62.97704
1991	Hindcast (Pop. based)	248.6478	403.1584	63.20375
1992	Hindcast (Pop. based)	248.9408	403.6334	63.27822
1993	Hindcast (Pop. based)	249.3191	404.2468	63.37438
1994	Hindcast (Pop. based)	250.0581	405.4449	63.56221
1995	Hindcast (Pop. based)	250.2331	405.7288	63.60672
1996	Hindcast (Pop. based)	250.2026	405.6793	63.59896
1997	Hindcast (Pop. based)	250.2160	405.7010	63.60237
1998	Hindcast (Pop. based)	250.0982	405.5100	63.57243

1999	Hindcast (Pop. based)	249.8823	405.1599	63.51754
2000	Hindcast (Pop. based)	249.3712	404.3312	63.38762
2001	Hindcast (Pop. based)	248.8597	403.5019	63.25761
2002	Hindcast (Pop. based)	248.2055	402.4411	63.09130
2003	Hindcast (Pop. based)	247.7323	401.6739	62.97104
2004	Hindcast (Pop. based)	247.4519	401.2193	62.89977
2005	Hindcast (Pop. based)	247.2849	400.9485	62.85732
2006	Hindcast (Pop. based)	247.0856	400.6253	62.80664
2007	Hindcast (Pop. based)	246.9826	400.4583	62.78046
2008	Hindcast (Pop. based)	247.2078	400.8234	62.83770
2009	Adjusted survey weights	164.5192	176.8699	71.14816
2010	Adjusted survey weights	289.7978	322.1593	100.22615
2011	Adjusted survey weights	201.0207	271.7714	25.65865
2012	Adjusted survey weights	198.0091	248.9475	114.50440
2013	Adjusted survey weights	276.0310	351.9106	74.23683
2014	Adjusted survey weights	487.2704	691.7426	73.63560
2015	Adjusted survey weights	252.0873	572.9438	26.09333
2016	Adjusted survey weights	111.4515	481.1934	71.65340
2017	Adjusted survey weights	162.1598	263.6850	16.83719

Total numbers

Weight and length of individual cod have been measured since mid 2016 and onwards.

$$TN_y = \frac{\sum_{y=16}^{ymax=18} weight_{adjusted,y} * 1000}{W_y/N_y}$$

where y is the year, $weight_{adjusted,y}$ is the weight in tonnes after being adjusted with the rekreaMultiplier, W_y is the weight of the cod registered in the REKREA survey in the year and N_y is the number of cod registered in the year.

Length and age

Due to limited data available, all data from the Otolith readings, are pooled to find the age length distribution in the years 2016 and 2017. Along with Swedish observations from 2012, 2016 and 2017.

Missing weights are estimated using a linear model based on the relation between log10(length) and log10(weight) for the cod in our data with both variables known. This model is also applied to the cod in the combined Danish and Swedish age length key in order to get more weight information in each year.

Age length key (ALK)

Using the same method as in (Jansen et al. 2008, 93–99), we find the current age length keys.

The age length key is an age distribution for each length from the sampling in REKREA and a mean weight at age and length. Due to the limited sampling we include Danish and Swedish catches from the recreational fishery, catches from the Danish commercial fishery and survey data from Danish and Swedish BITS cruises in the Sound.

With this amount of data, we still lack data for years 1980-1990 and 2001-2003. To compensate for this, we make an age length key based mean values on the years 1991-1994 and apply this to the years 1980-1990 and a age length key based on mean values on the years 1997-2000 and 2004-2008 which we apply to the years 2001-2003.

Age length distribution (ALD)

In the ALD, ALK numbers are converted to fraction normalized to 1 for each length in each year making the sum of each row 1.

$$Fraction_{Y,a,l} = \frac{N_{Y,a,l}}{N_{Y,l}}$$

Where $N_{a,l}$ is the number of cod in year Y at a certain age, a, and length, l, and N_l is the number of cod at the certain length.

CATON

We have the total catch from our adjustments of survey data and the hindcast going back to 1980. Releases are included in these numbers with a post release mortality rate of 11,2%.

$$CATON_{Y,a} = CATON_{Y,Total} * \frac{N_{Y,a}}{\sum_{a=1}^{a=12} N_{Y,a}}$$

Where$CATON_{Y,Total}$ is the combined harvest and 11,2% of releases per year, $N_{Y,a}$ is the number of cod in the year Y with the age a found in the age length keys each year.

CANUM

CANUM for the year Y and age a is found as:

$$CANUM_{Y,a} = \frac{CATON_Y}{\left(\sum_{l_{min}}^{l_{max}} meanWeight_l * Fraction_l\right)} * Fraction_{Y,a,l} * Fraction_l$$

Where $CATON_Y$ is the catch in tons in a year, $meanWeight_l$ is the mean weight at a given length, $Fraction_l$ is the fraction of a fish being at a given length, $Fraction_{Y,a,l}$is the fraction of cod in a year, at a certain age and length.

WECA tables

The mean weight at age is found for the sampled cod in Danish and Swedish catches from the recreational fishery and catches from the Danish commercial fishery.

$$meanAge_{Y,a} = \frac{\sum weight_{Y,a}}{N_{Y,a}}$$

Where y is the year of sampling, $weight_{age,y}$ is the weight of age read cod in a sampling year and $N_{age,y}$ is the number of age read cod at a certain age in a sampling year.

Assumptions

Deliverable	**Area**	**Assumptions**
CATON	SD22	Same relationship between DST and REKREA catch data as in SD23. Reatio between DST and REKREA charter/private boat catch estimates is the same for landbased/coastal fishing. Same effort, i.e. fraction of the Danish population fishing from 1980 to 2009.
CATON	SD23	Ratio between DST and REKREA catches is based on effort and catch data from 2016 – 2018 (8 quarters). Same ratio (mean) is assumed and applied back to 2009.

		Reatio between DST and REKREA charter/private boat catch estimates is the same for landbased/coastal fishing. Same fraction of the Danish population fishing from 1980 to 2009.
CATON	SD24	Same relationship between DST and REKREA catch data as in SD23. Reatio between DST and REKREA charter/private boat catch estimates is the same for landbased/coastal fishing. Same effort i.e. fraction of the Danish population fishing from 1980 to 2009.
CANUM	SD22	Same age at length for commercial, BITS and recreational caught cod Length distribution and age at length for Danish and German catches is the same
CANUM	SD23	Charterboats and private boat targets and catches the same sizes of cod Same age at length for commercial, BITS and recreational caught cod Same length distribution for Danish and Swedish caught cod
CANUM	SD24	Same age at length for commercial, BITS and recreational caught cod Length distribution and age at length for Danish and German caught cod is the same

References

Jansen, Teunis, Henrik Degel, Kirsten Birch Håkansson, Josefine Egekvist, Jørgen Dalskov, and Fritz Köster. 2008. "FishFrame 5.0: A Web Based Datawarehouse Application for Management, Access and Integration of Fisheries and Stock Assessment Data." In *ICES Council Meeting 2008*, 26:1–137.

Sparrevohn, Claus Reedtz, and Marie Storr-Paulsen. 2010. *Åle- Og Torskefangst Ved Rekreativt Fiskeri I Danmark: Undersøgelsesdesign Og Fangster I 2009*. DTU Aqua. Institut for Akvatiske Ressourcer.

———. 2012. “Using Interview-Based Recall Surveys to Estimate Cod Gadus Morhua and Eel Anguilla Anguilla Harvest in Danish Recreational Fishing.” *I C E S Journal of Marine Science* 69 (2). Oxford University Press: 323–30. doi:10.1093/icesjms/fss005.

Statistik, Danmarks. 2018. “Danmark Statistik–Statistikbanken [Statbank Denmark].” *URL: Http://Statistikbanken.dk/Statbank5a/Default.asp?w=1920*.

WD 16 - Reconstructing Swedish cod catch numbers at age for tour-boat and private data in SD23

Andreas Sundelöf[1], Hege Sande[1], Esha Mohamed[1]

[1]Department of Aquatic resources
Swedish University of Agricultural Sciences
Turistgatan 5, 453 30 Lysekil, Sweden

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Contents

Intro

Swedish recreational fishing targeting cod occurs mainly from tour-boats and private boats. Catch estimates come from logbooks of total catch (2011-2017) and from marina sampling and on-board sampling of tour-boats since 2017. Thus sizes and ages are rare information and logbooks do not cover the entire time series handled by the assessment model SAM (1994-2017).

To bridge this obstacle to enter Swedish RecFish data into SAM a learning algorithm has been used to reconstruct total catches from tour-boats. Since this data lacks sizes SLU has used available size data from 2012, 2013 and 2017 to fit a selectivity of the tourboat catch data to sizes from the BITS survey of the same years. The BITS time series is then used to reconstrict size data through the time series which allows for age conversions such that a CANUM matrix can be calculated for the tour-boat data.

Fishing from private boats are added to the tour-boat data series.

Method

Annual catches

Face value annual catch of the tour boat fleet was used 2011-2017

For 2017 also estimated catches from private fishing was added to the tour boat catches.

Catch reconstructed 2010 - 1980

Catch estimates of tour boats from 2011-2017 have been reconstructed back to 1994 through a learning algorithm using auxiliary time-series data (temperatures, wind and recalled effort) described in Appendix 1 (Mohammed and Sundelöf 2018).

- Features that were available for model training were Year, month, temperature for Falsterbo, wind for Falsterbo and Helsingborg and catches from a retrospective study. In the retrospective study, a total of nine captains from different access sites were interviewed about the effort and its trend back to 1994 and an average proportion of the catches contributed in each month. The effort and the proportion were multiplied to obtain a relative absolute effort for each month between 1994 and 2017.
- Catch was reconstructed by month for 1993-2010 and then aggregated to yearly catches.

- Previous to 1994 there was no estimated effort from tour boats. In order to reconstruct estimated annual catches 1980-1993 the average of the reconstructed catches 1994-1998 was used.

Sizes, ALK

Size

- Face value size composition was used for 2012, 2013 and 2017. For other years we used the average observed size composition of combined Swedish and Danish length measurements from on board tour boat sampling.
- Swedish and Danish data on fish individuals have been combined.

ALK

- Growth was dependent on age in the BITS series. Converting sizes to ages was performed using an age-length key for each individual year. R package {Fishmethods} was used to calculate ALK type 2, i.e. proportions-at-age per size.
- ALK and WECA were calculated from BITS/commercial data from the Swedish and Danish sampling programs (years?)

Releases

Releases were not included in the catch.

A sigmoid selectivity function was fitted to calculate released fish "discard" depending on the length of the fish from Swedish on-board tour boat sampling in 2017. The function was fitted to observations of lengths of released fish and fitted by minimizing sum of squared residuals by adjusting parameters 1 and 2 in the following expression:

prob_retention = K*par1*exp(par2*Length)/K+par1*(exp(par2*Length)-1)

The final parameters were

K	1
par1	6.65E-05
par2	0.203293

Catch and release survival is regarded high in this cod fishery (Welterbach et al 2017). Hence returned fish can be estimated from and reconstructed from 2017. We do not provide an estimate of surviving returned fish with an anticipated large variance. We assume that releases survive.

Estimation of releases is possible from back calculation of the probability of retention of a fish of a certain length given a full length distribution of captured fish.

CANUM

Catch numbers at age in the Swedish and Danish recreational fishery for WBC were computed through combined ALK and WECA by DTU Aqua.

WD 17 - Maturity of Western Baltic cod

Data

Proportions mature at age for WB cod are based on BITS Q1 survey. Only data from SDs 22-23 are used, as SD24 consists of a mix of eastern and western Baltic cod.

SDs are defined based on coordinates given for individual fish data in DATRAS:

SD 22: ShootLat>53.5000 & ShootLat<=56.0000 & ShootLong>9.5000& ShootLong]<=12.0000
SD 23: ShootLat>=55.5000& ShootLat<=56.0000& ShootLong>12.0000& ShootLong<13.0000

German and Danish data are pooled for the analyses. The number of individual fish for which maturity data are available in each year is shown in Table 1.

Table 1. Number of WB cod for which maturity data are available from BITS survey (SD 22-23, Q1).

Year	Number of fish
1996	307
1997	369
1998	256
1999	838
2000	336
2001	715
2002	705
2003	400
2004	959
2005	844
2006	795
2007	929
2008	520
2009	411
2010	727
2011	1008
2012	623
2013	567
2014	531
2015	721
2016	363
2017	609
2018	792
Grand Total	**14325**

Interpretation of maturity stages

The maturity stages by country uploaded to DATRAS database were interpreted as follows:

Country	Years	Stages in DATRAS	Interpretation
DEN	1996-2002 2004-2017	1-5 61-66	1,5: non-spawner 2-4: spawner 61, 65, 66: non-spawner 62-64 : spawner
DE	1999-2009 2010-2017	1-5 I-IX	1,5: non-spawner 2-4: spawner I,II, IX: non-spawner III-VIII: spawner

Results

The calculation of proportion of spawners in the population was scrutinized to ensure consistent treatment of data in the entire time series. This resulted in some revisions to the time series used by WGBFAS previously, though the overall calculation procedure is unchanged and the revisions do not affect the overall dynamics in maturation. Part of the revision is due to the area now being defined by coordinates, while it previously was based on statistical rectangles. Figure 1 shows the comparison on annual proportions of spawners in the new revised time series compared to the previous one used in WGBFAS.

In stock assessment, proportion of spawners has been used as 3 years running mean, e.g. the values for 2017 used in assessment were based on the average of the estimates for 2015-2017. 3 years means have been applied due to a relatively large variability in the data. As the 3-year running mean still contains considerable inter-annual variability in the data to be used in stock assessment, WKBALTCOD2 data meeting agreed to use a 5-year running mean instead, that captures the long term trends but reduces inter-annual fluctuations (Figure 2).

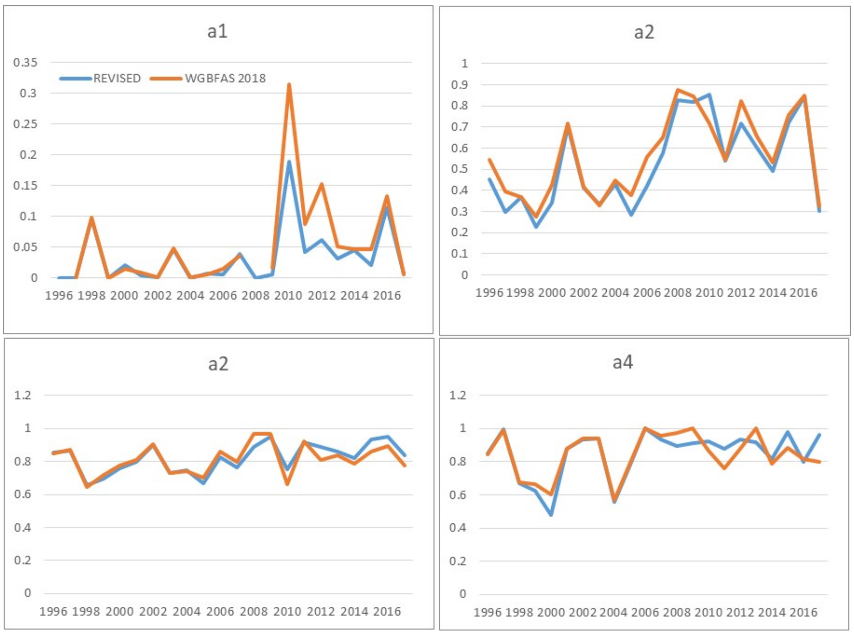


Figure 1. Comparison of revised annual values for maturity at age with the values used for last year's assessment.

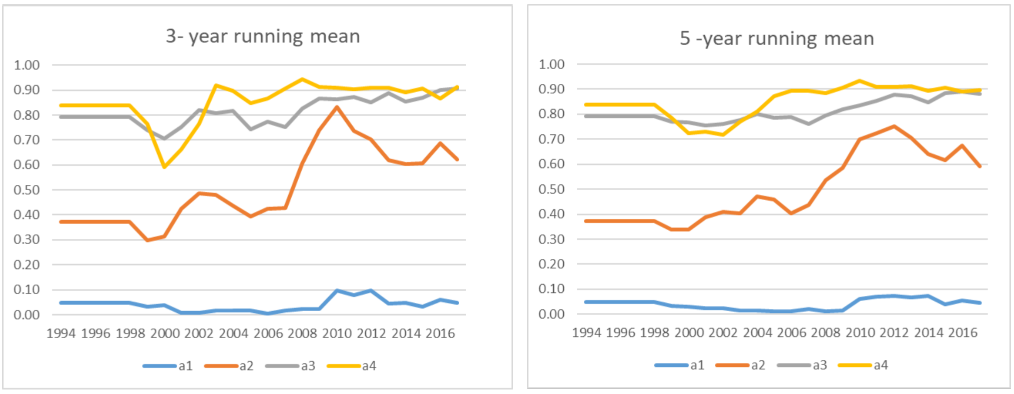


Figure 2. 3-year running mean proportion mature at age compared to a 5-year running mean.

WD18 - Independent western Baltic cod recruitment index: juvenile cod data from commercial pound nets

Kate McQueen, Uwe Krumme

Thünen Institute of Baltic Sea Fisheries, Alter Hafen Süd 2, 18069 Rostock, Germany

Introduction

The recruitment index for western Baltic cod is currently estimated from data collected during the Baltic International Bottom Trawl Surveys (BITS) in quarter 1 (Q1) and quarter 4 (Q4). The trawl survey covers areas in the western Baltic Sea deeper than10m (ICES, 2017). However, juvenile cod are reported to inhabit shallow inshore waters (Pihl and Ulmestrand, 1993) and may preferentially occupy shallow-water vegetated habitats (Freitas et al., 2016) which are not adequately covered by the BITS. Therefore, an independent sampling programme which targets juvenile cod in their preferred habitat could prove useful in improving the accuracy of estimations of cod year-class strength.

Additionally, the scientific surveys are often criticised by fishers, raising justified arguments (e.g. snapshot data only) and unjustified criticisms (e.g. fishing where there is no fish) and rendering the communication of scientific advice to fishers a difficult task. Therefore, considering direct, scientifically validated information from the commercial fishery may facilitate the communication with these key stakeholders and increase their understanding of and confidence in the data used in fish stock assessment.

We analysed data from a specialised sampling programme targeting juvenile cod captured in commercial pound nets set in shallow waters off the coast of Fehmarn in Germany to explore the potential of this data as an additional, independent source for estimating recruitment of cod in the western Baltic Sea.

Methods and Results

Pound net sampling

Cod samples were provided by two full-time, commercial pound net fishers who operate from the island of Fehmarn in Germany. The pound net fishers each maintain 3-5 stationary, uncovered pound nets year round in shallow waters (<5m) close to the coast (Figure 1). The nets are set perpendicularly to the coast in seagrass dominated substratum, and span the entire water column. Buoys at the head line and weights at the lead line force fish to enter the catch chamber (mesh size 12mm) at the seaward end. The catch chamber is stretched by ropes attached to fixed pillars (Figure

2). The pound net fishers target eel, but undersized cod are also retained in the catch chamber, when they migrate between daytime resting sites in deeper waters and nightly feeding sites in structured shallow water habitats (Burrows et al., 1994).

Cod samples were collected from the pound net fishers in 2011, and during 2013-2017. In 2011 samples of undersized cod from the pound nets were collected between 1 and 5 times a month during August to December as part of an observer sampling scheme. During 2013-2017 the catch chamber was emptied by the fishers every 1-17 days (mean 2.3 days) depending on the weather conditions during the fishing seasons (April-June and September-December). From 2014 onwards, efforts were made to collect samples year round, in addition to the fishing seasons.

Samples (average sample weight: 3kg) of small cod (<38cm in length) were collected and weighed, and the weight of the entire catch including the sub-sample was estimated by the fisher and recorded. In October 2013 and November 2014, a small sub-sample (2-5%) of undersized cod were measured live, and then used for another experiment. From 2014 onwards, sub-samples (22%) of undersized cod were measured, tagged and re-released as part of an ongoing age validation study (Krumme et al., unpublished data). The remainder of undersized cod used for the recruitment index were frozen immediately (-20°C) after landing at the port (Burgstaaken, Fehmarn). Cod from the frozen samples were later defrosted, measured and processed at the Thuenen Institute of Baltic Sea Fisheries (TI-OF).

Standardisation of length frequency data from pound net samples

To standardise the length frequencies for effort and total size of catch, the ratio between the total weight of the catch (estimated by the fisher) and the sample weight was used to raise the length frequency to the total size of the catch. Since the undersized cod samples usually came from a combination of the pound nets maintained by each fisher, the number of cod per length class estimated for the total catch was divided by the total number of nets which contributed to the catch, and by the total number of days soaking time of the nets, to calculate the average number of cod per length class per net per day.

In 2011, data on soaking times of the nets before sampling and number of nets which contributed to the catch were not recorded. The total weight of the catch was not estimated; instead observers measured all individuals within the catch. To allow this data to be used in analysis, so that the entire year did not have to be excluded, the number of nets and days soaking were set to 1 for analysis.

The cod held alive for tagging were not included in the estimated total weight of catch. They were treated separately from the frozen samples gathered during the same period, and so were given a raising factor of 1. Generally, the larger individuals were selected for tagging if they were in good condition (mean length of tagged individuals: 25cm, mean length of samples without tagged individuals: 22cm). If they were excluded from analysis because of uncertainty surrounding the appropriate raising factor, the size distribution of the samples would be biased (Figure 3). During 2014-2017 8386 cod were removed for tagging, compared to 29882 which were frozen for length frequency samples (i.e. tagging data accounts for 21.8% of the total sample size from 2014 onwards).

In October 2013 the data for the individuals measured live for another experiment were not used in subsequent analysis because of a lack of information required for the raising factor. In November 2014 all necessary data were available to include these individuals in subsequent analyses.

Monthly averages were estimated from the standardised abundances of cod caught per length class, per length per day in each sample. Averages were not weighted by fishing effort; each sample provided had equal weight.

Identification of cohorts within length frequencies

There were two clearly defined modes between 10 and 38cm in the monthly standardised length-frequencies (Figure 4). These were most easily identified during September-December, when cod were most abundant in the pound net catches. By following the progression of length frequency modes, it can be reasonably assumed that the mode which can be tracked from April until the end of the year, with average length increasing from around 20-25cm to 30-35cm, represents the age-1 cohort. A new cohort appears in the length frequency in September. The average length of this cohort increases from around 12cm to 18cm by December. This is likely to be the young-of-the-year fish, spawned in spring and grown large enough to be retained in the pound nets by autumn, when they begin making use of the shallow water habitat where the pound nets are set.

Analysis of otoliths to confirm age interpretation

To investigate the validity of assigning cohorts based on the length frequency modes, the otoliths from undersized cod sampled between 2013-2016 were examined.

The otoliths from sampled cod were removed and sectioned using standard procedures. The right otolith from each sampled cod was embedded in GTS polyester casting resin (Voss Chemie, 35-40% Styrol) with MEKP hardener. The otoliths were thin-sectioned (thickness 0.5mm) through the core using an ATM Brilliant 250 bone saw. Images of each individual otolith were taken using transmitted light and a light microscope, with the Zen Blue software (Carl Zeiss). Under these conditions the opaque zones appeared darker and the translucent zone (TZ) lighter than surrounding material.

The otoliths collected during 2013-2016 were classified based on the number of completed TZs within the otolith and the edge type of the otolith (i.e. the classification of zone forming at the marginal edge of the otoliths). Each individual was assigned an "edge-zone category" as follows: an otolith with 0 completed TZs and a translucent edge was classified as 0t, an otolith with 1 completed TZs and an opaque edge was classified as 1o and so on. The numbers of edge-zone categories classified per length class were plotted (Figure 5) and the proportions of edge-zone categories per length class were overlaid onto the standardised length frequencies (Figure 6) (McQueen et al., 2018).

An age-reading table was consulted to assign ages to individuals based on their month of capture and edge-zone category (Table 1, McQueen et al., 2018). To conform to standard Baltic cod age reading methods, the birthday of all cod was set to 1st January.

The otolith assigned ages confirmed the assumption that the first length frequency mode visible in the samples between September and December represented the age-0 cod, and the second mode represented the age-1 cod (Figure 7).

Modal decomposition of length frequencies

Parameter estimates for the overlapping normal distributions which best fit the standardised length frequencies in September to December were estimated using a combination of Newton-type method and EM algorithm, applied through the R package "mixdist" (McDonald, 2018). Two normal distributions were fit to each monthly average distribution, except September 2011, 2013, 2015 and 2017, and October and November 2017, when only the age-1 cohort was visible in the samples. The monthly average abundance per cohort, per net, per day was extracted by summing the joint distributions. Additionally, the sample data for October-December and September-November were combined, and the average abundances per net per day for age 0 cod during October-December and for age 1 cod during September-November were estimated in the same way. These combinations of months were considered the most representative for each age class, given that in some years the age-0 cohort was not yet clearly visible in the samples in September, and in December the abundances of age-1 cod in the pound net samples were reduced relative to the previous months. It was not possible to detect the age-0 peak in the 2017 data, thus the average pound net abundance for age-0 cod in 2017 over the three month period was set to 0.

The abundance indices for month combinations described above were calculated by pooling all the standardised samples together before averaging (Abundance Index 1, Table 2). However, this method does not correct for temporal variations in sampling effort between years (Figure 8).

Therefore, for age 0 abundances between October to December, the monthly averages were also averaged together to produce a combined abundance index for which each month essentially had equal weight (Abundance Index 2, Table 2). These results were also provided for age 0 as an alternative combined month abundance index, and demonstrated that the two averaging approaches produced slightly different results (Table 2).

Variance was also estimated for these two abundance indices. For abundance index 1, variance in number of fish caught per net per day between samples collected during October to December was estimated for each length class. Variance of each length class was multiplied by the square of the conditional probability of being in the age-0 cohort (from the fitted normal distributions). The probability corrected variances for each length were summed to calculate the overall variance of the abundance index of the age-0 cohort. For abundance index 2, the variance between the total abundance indices of each month within the year was calculated.

Comparison of pound net and BITS abundances

The average abundances of age-0 and age-1 cod estimated from the pound net samples were compared to the BITS estimated abundances. The combinations of BITS indices and pound net samples which were compared using linear regressions are listed in Table 3. Each reasonable combination of BITS sample and month of pound net sampling was tested, as it was unknown which month of pound net sampling would provide the best indication of juvenile cod abundance. Model fits were assessed through comparisons of R^2.

The highest R^2 values were found for models which compared the age-0 cohort abundances estimated from the pound net samples between October and December (either separately or pooled (Abundance Index 1)), to the age-0 abundance in BITS Q4, or the age-1 abundances in BITS Q1 (Figure 9). However, it is important to note that these relationships appear to be driven by the extremely large year-class in 2016, as the data for 2016 have very high leverage within the regression models (extremely high Cook's distance, significantly higher than all other data points). When the analysis was repeated, excluding data from 2016, the R^2 was considerably reduced (Table 3).

Summary

The standardised length frequency samples of juvenile cod collected from the pound nets set in shallow waters around Fehmarn clearly represent the progression of age-0 and age-1 cohorts. The abundances of juvenile cod estimated from the pound net sampling follow a similar pattern to the abundances estimated from BITS, but the trends do not match exactly. However, the trawl survey and pound nets sample juvenile cod from different habitats, so a perfect match between the independent datasets should perhaps not be expected. The key advantage of the pound nets is that they sample juvenile cod from their preferred shallow, vegetated habitats. Therefore, the data can be considered to be more representative of juvenile cod abundances in this region between October-December than trawl sampling conducted in deeper regions.

Moreover, the pound nets operate from September to December and are regularly emptied so that the cohort strength is sampled with high temporal resolution throughout the period in which juvenile cod use the shallow water habitat. This gives an important independent and additional source of information.

While the abundance changes in the pound nets clearly confirmed the strong 2016 and the weak 2015 and 2017 year classes detected by BITS, the high R^2 are also strongly driven by these extremes. However, the close correlations during these three extreme years also shows that the pound net index reliably detects minima and maxima in year class strength, an important prerequisite for model predictions.

The pound net index represents a scientifically validated independent data set originating from the commercial fishery. Since 2015 the results from these data have enormously helped in confirming the extremes detected by BITS, explaining the western Baltic cod population dynamics to the fishery and other stakeholders. Ultimately, they have increased the trust of both scientists and fishers in the data and in the model outcomes and the advice.

References

Burrows, M.T., Gibson, R.N., Robb, L., Comely, C.A., 1994. Temporal patterns of movement in juvenile flatfishes and their predators: underwater television observations. J. Exp. Mar. Biol. Ecol. 177, 251–268.

Freitas, C., Olsen, E.M., Knutsen, H., Albretsen, J., Moland, E., 2016. Temperature-associated habitat selection in a cold-water marine fish. J. Anim. Ecol. 85, 628–637. https://doi.org/10.1111/1365-2656.12458

ICES, 2017. Manual for the Baltic International Trawl Surveys (BITS). Ser. ICES Surv. Protoc. SISP 7 - BITS 95. https://doi.org/10.17895/ices.pub.2883

McDonald, P., 2018. mixdist: Finite Mixture Distribution Models. R package version 0.5-5. https://CRAN.R-project.org/package=mixdist

McQueen, K., Hrabowski, J., Krumme, U., 2018. Age validation of juvenile cod in the western Baltic Sea. ICES J. Mar. Sci. DOI:10.1093/icesjms/fsy175

Pihl, L., Ulmestrand, M., 1993. Migration pattern of juvenile cod (Gadus morhua) on the Swedish west coast. ICES J. Mar. Sci. 50, 63–70.

Tables

Table 1: Matrix of assigned age based on month and edge-zone category, using the classification system outlined in McQueen et al., 2018 (under review). UC indicates unclassifiable combinations of month and edge-zone category (due to the potential for misclassification). Cells filled with a dash (-) indicate combinations of month and edge-zone category which are very unlikely to occur following the current understanding of patterns in TZ formation, and which were very rarely observed in the samples (Figure 6).

	Edge-zone category					
Month	**0o**	**0t**	**1o**	**1t**	**2o**	**2t**
January	-	-	1	-	2	-
February	-	-	1	-	2	-
March	-	-	1	-	2	-
April	-	-	1	-	2	-
May	-	-	1	-	2	-
June	-	-	1	1	2	2
July	-	-	1	1	2	2
August	0	0	UC	1	UC	2
September	0	0	0	1	1	2
October	0	0	0	1	1	2
November	0	0	0	1	1	2
December	0	0	0	1	1	2

Table 2: Comparison of the mean abundance indices of age 0 cod captured in pound nets between October and December used in the linear regressions with BITS data (Abundance Index 1, Table 3), and abundance indices calculated from the monthly averages, so that each month has equal weight to account for variation in sampling effort (Abundance Index 2).

Year	Month combination	Age	Abundance Index 1: weighted by distribution of samples	Abundance Index 2: each month has equal weight
2011	Oct - Dec	0	15.9	20.7
2013	Oct - Dec	0	12.9	16.9
2014	Oct - Dec	0	23.1	25.6
2015	Oct - Dec	0	2.5	4.3
2016	Oct - Dec	0	187.0	164.2
2017	Oct - Dec	0	0	0.4

Table 3: Results of linear regressions between BITS data and pound net data. The year column indicates whether pound net samples and BITS data were collected during the same year (same) or whether the BITS data were collected the year after the pound net data (next). For analysis using the average abundance of age-0 cod in the pound nets during October-December (abundance index 1), analyses were run either with data from 2017 excluded, or with average abundance in the pound nets during 2017 set to 0, as no mode was visible for this cohort in the length frequency. As data from 2016 were found to disproportionately influence some of the linear regressions, analyses were repeated with 2016 removed, for comparison. Analyses with a given year of data excluded are indicated by the "year excluded" column. "NA" in the R^2 column indicates tests which could not be run due to lack of data. Tests which resulted in an R^2 0f >0.9 are shaded yellow.

BITS data			Pound net data		Year excluded	Adjusted R^2
Age	Quarter	Year	Age	Month		
0	4	same	0	Sep		NA
0	4	same	0	Oct		0.9718
0	4	same	0	Nov		0.9688
0	4	same	0	Dec		0.9405
0	4	same	0	Oct-Dec	2017	0.9839
0	4	same	0	Oct-Dec		0.9855
0	4	same	0	Oct-Dec	2016	0.5235
1	1	next	0	Sep		NA
1	1	next	0	Oct		0.9564
1	1	next	0	Nov		0.959
1	1	next	0	Dec		0.9434
1	1	next	0	Oct-Dec	2017	0.9717
1	1	next	0	Oct-Dec		0.9751
1	1	next	0	Oct-Dec	2016	0.3485
1	1	same	1	Sep		0.5566
1	1	same	1	Oct		0.7851
1	1	same	1	Nov		0.4862
1	1	same	1	Dec		-0.2365
1	1	same	1	Sep-Nov		0.5253
1	4	same	1	Sep		0.1022
1	4	same	1	Oct		0.3781
1	4	same	1	Nov		0.06
1	4	same	1	Dec		-0.3281
	4	same	1	Sep-Nov		0.3226
2	1	next	1	Sep		0.1017
2	1	next	1	Oct		0.3773
2	1	next	1	Nov		0.02524
2	1	next	1	Dec		-0.3172
2	1	next	1	Sep-Nov		0.067

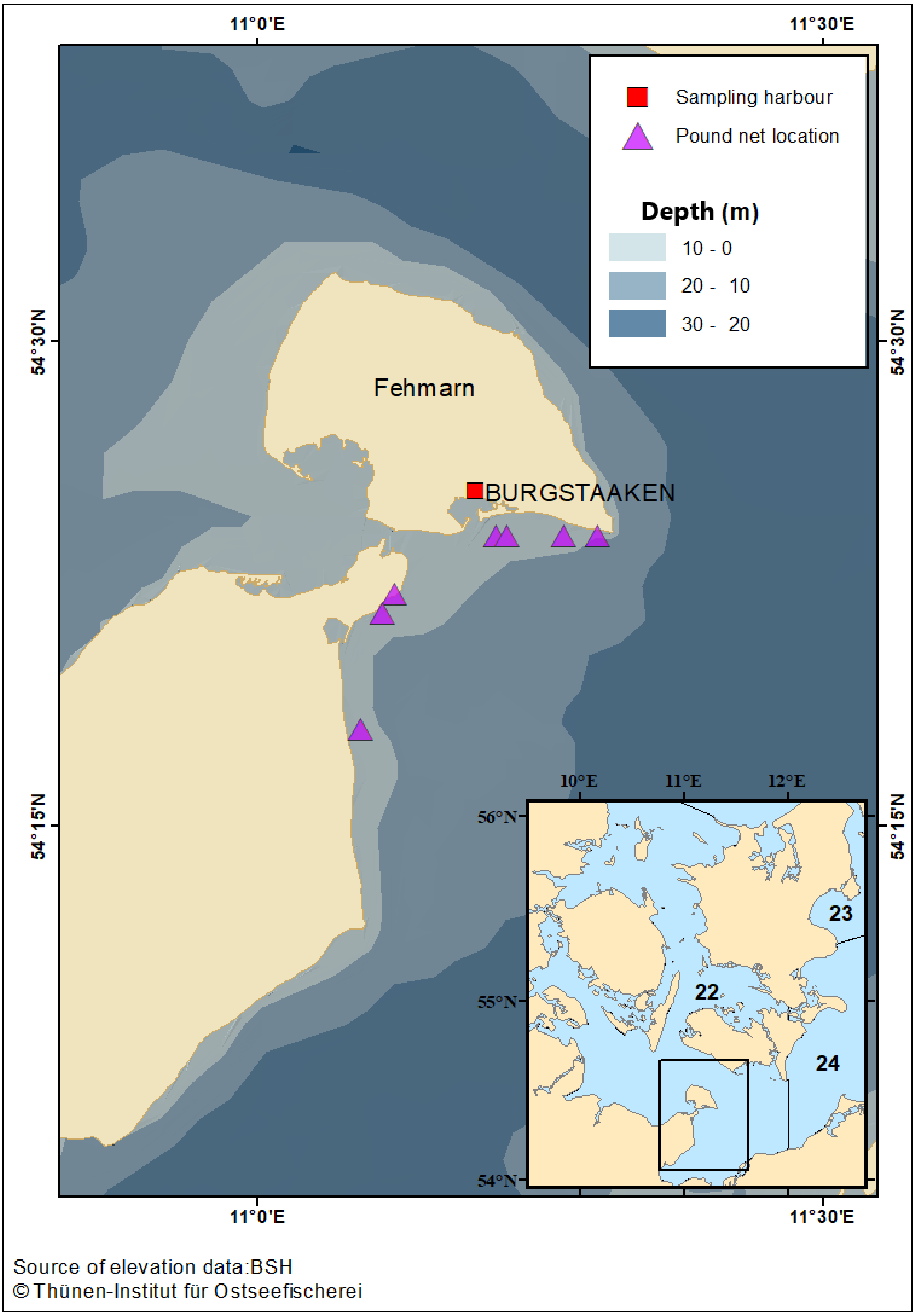
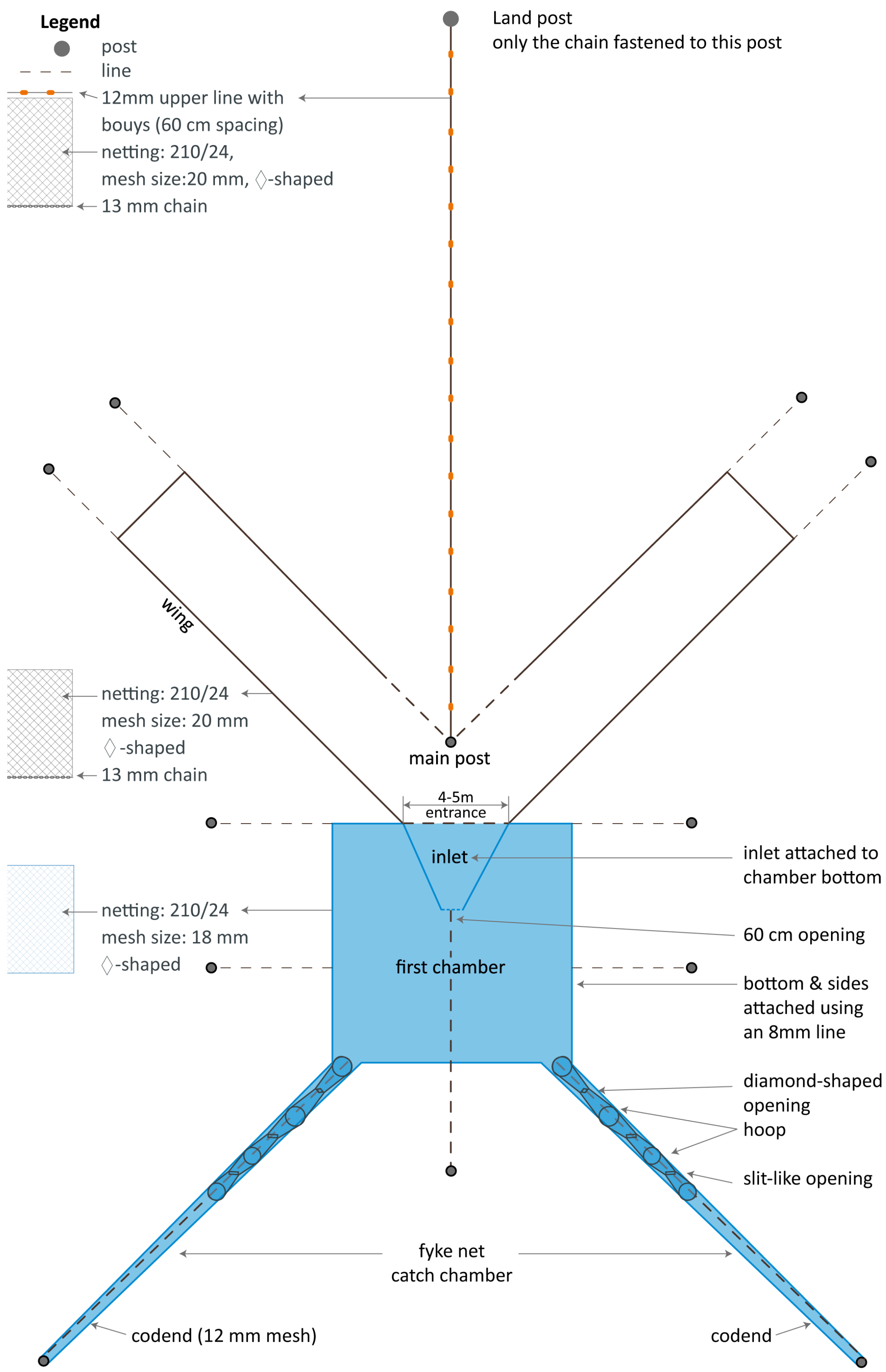


Figure 1: Location of pound nets off the coast of Fehmarn, from which samples were collected.



Legend
post
line
12mm upper line with
bouys (60 cm spacing)
netting: 210/24,
mesh size:20 mm, ◊-shaped
13 mm chain
Land post
only the chain fastened to this post
wing
netting: 210/24
mesh size: 20 mm
◊ -shaped
13 mm chain
main post
4-5m
entrance
inlet
inlet attached to
chamber bottom
netting: 210/24
mesh size: 18 mm
◊-shaped
60 cm opening
first chamber
bottom & sides
attached using
an 8mm line
diamond-shaped
opening
hoop
slit-like opening
fyke net
catch chamber
codend (12 mm mesh)
codend

Figure 2: Upper figure: Top-down perspective of a commercial pound net in Fehmarn used to trap the sampled juvenile cod. Original drawing: E. Pahlke, digital redrawing: Thünen Institute of Baltic Sea fisheries (A. Schütz);lower figure: a commercial pound net near the mouth of Burgstaaken harbour, Fehmarn; from front to back: first chamber, wings and guiding net or leader (with white buoys), beach (photo: U. Krumme).

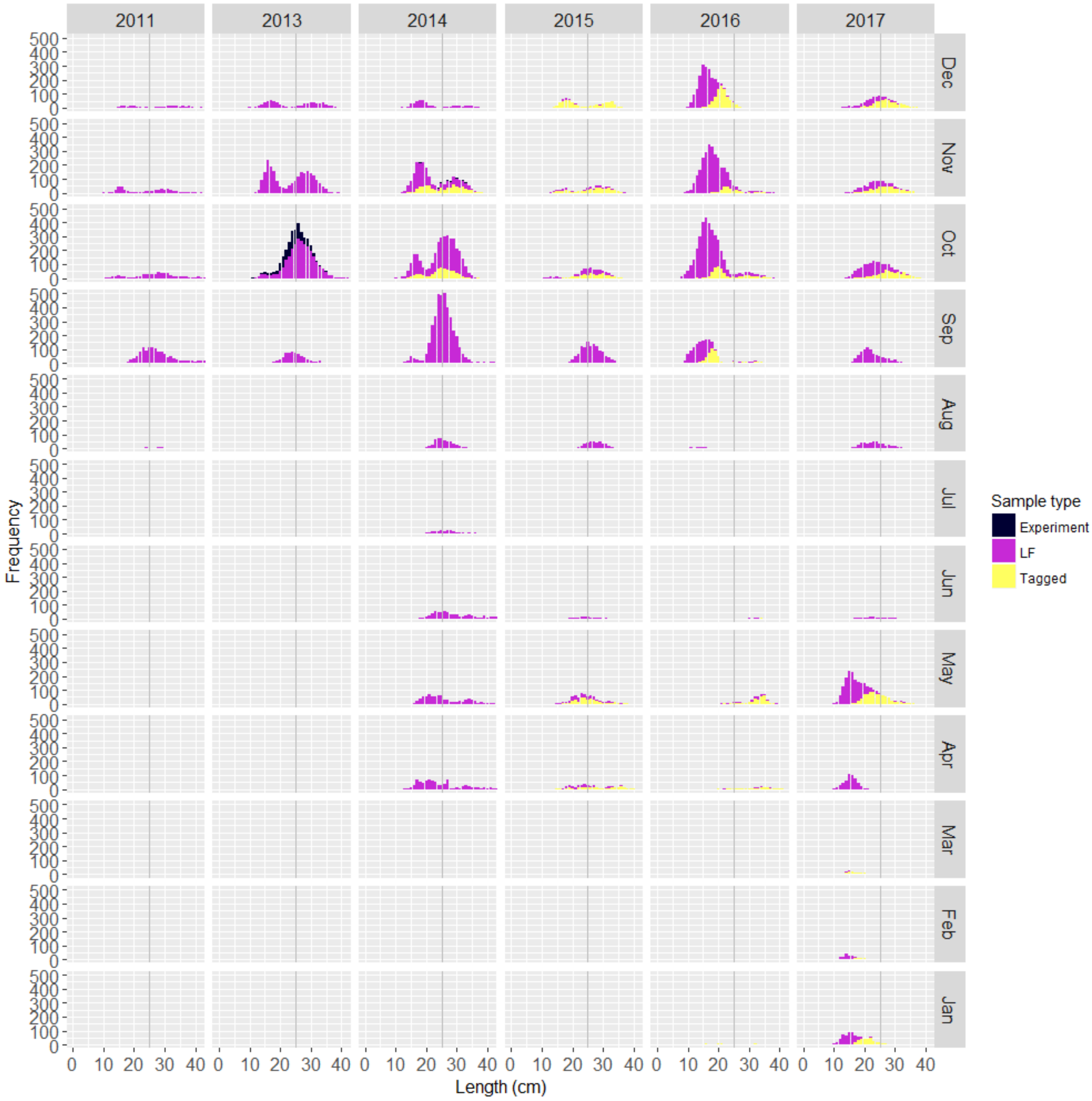


Figure 3: Length frequency distributions, showing true sample sizes each month and each year, without correcting for fishing effort or total size of catch. “LF” samples are the cod frozen immediately and measured to provide length frequency data; “Experiment” samples are those measured live and then used for another experiment; “Tagged” individuals are those measured live, then tagged and released as part of an ongoing age validation experiment.

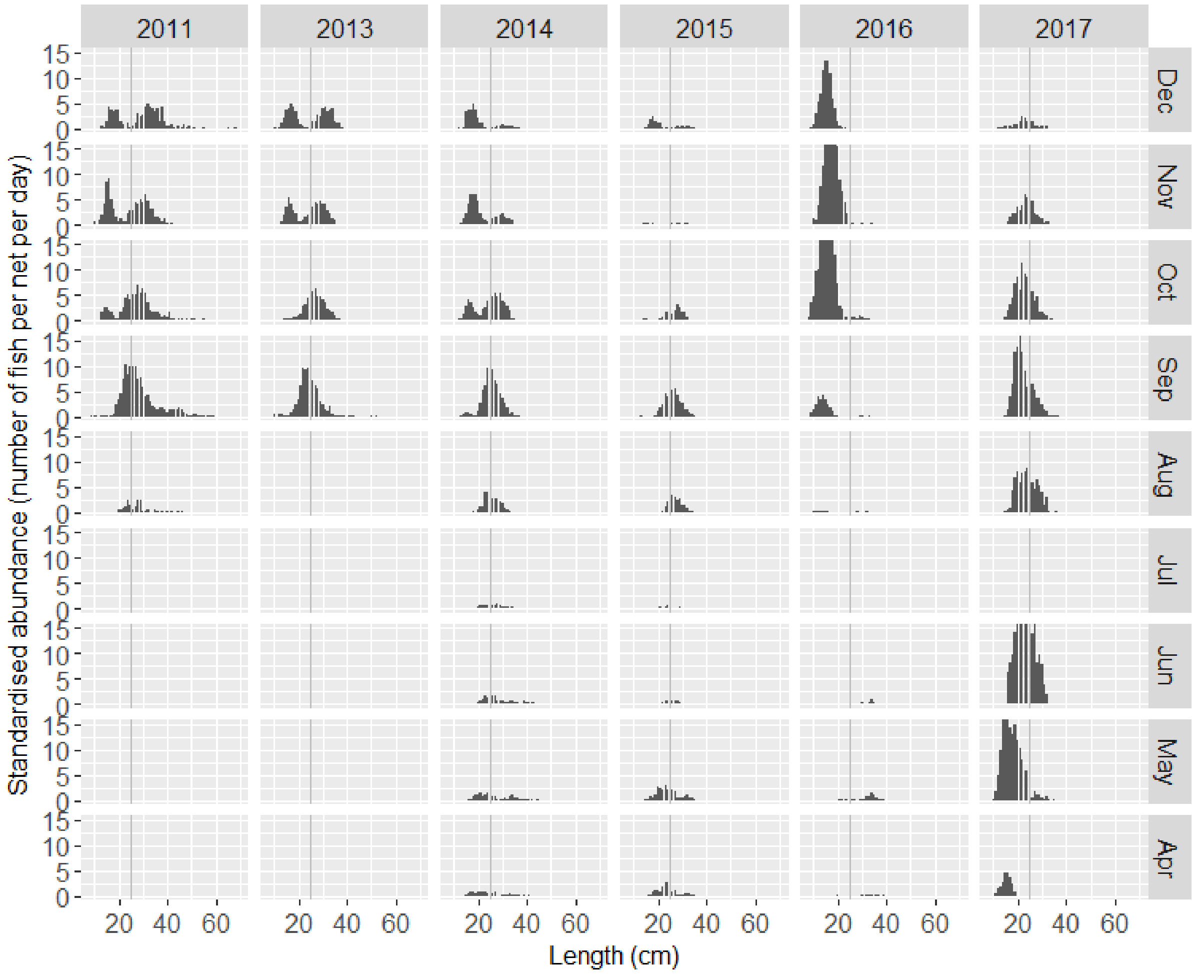


Figure 4: Length frequencies, showing the average catch per month, standardised for fishing effort and total size of catch. The y-axes are truncated between 0 and 15, due to high frequencies in 2016 Oct-Dec.

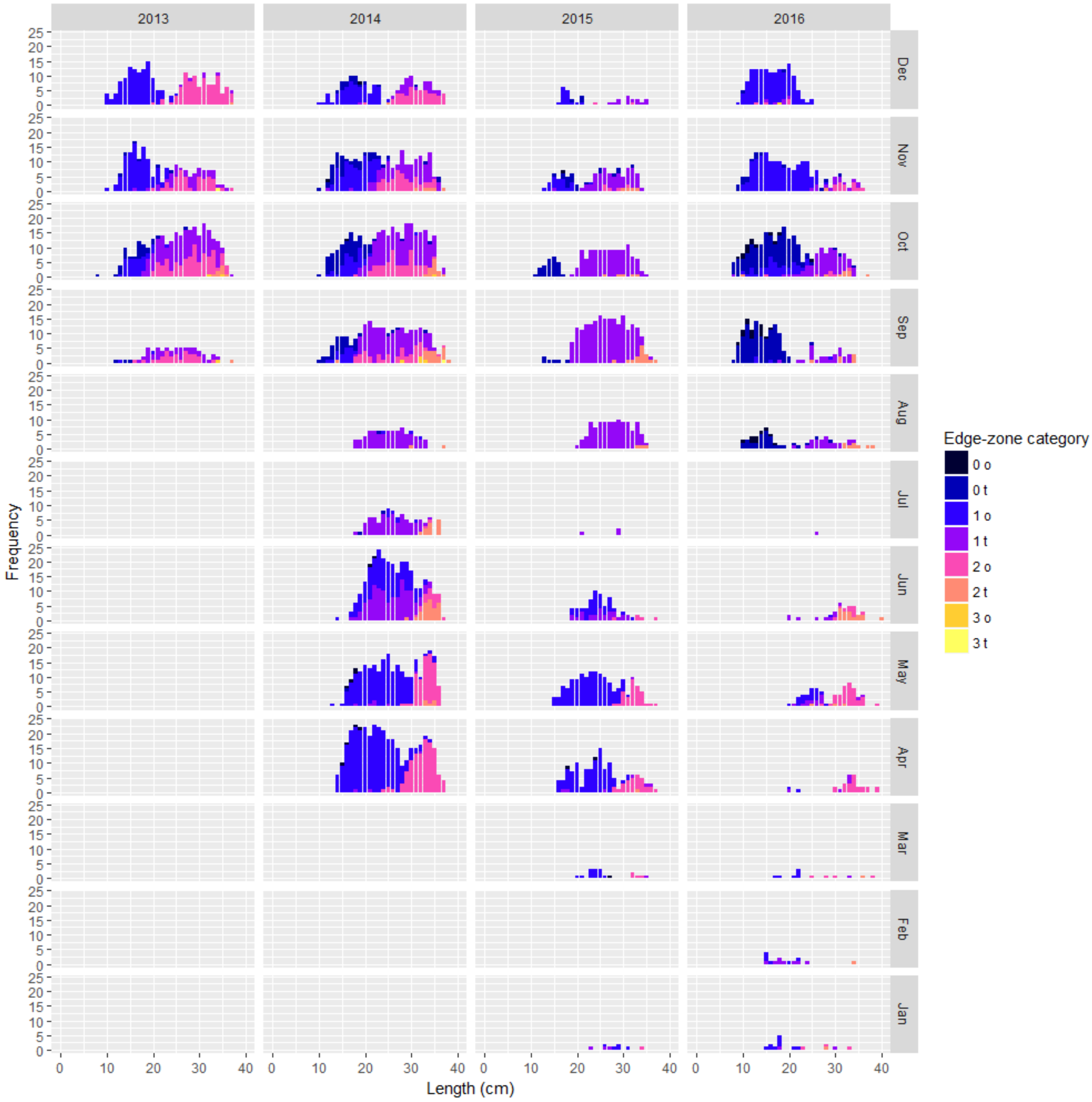


Figure 5: Edge-zone categories assigned to every otolith analysed, illustrating the frequency of edge-zone categories per length class for each sampling month and year.

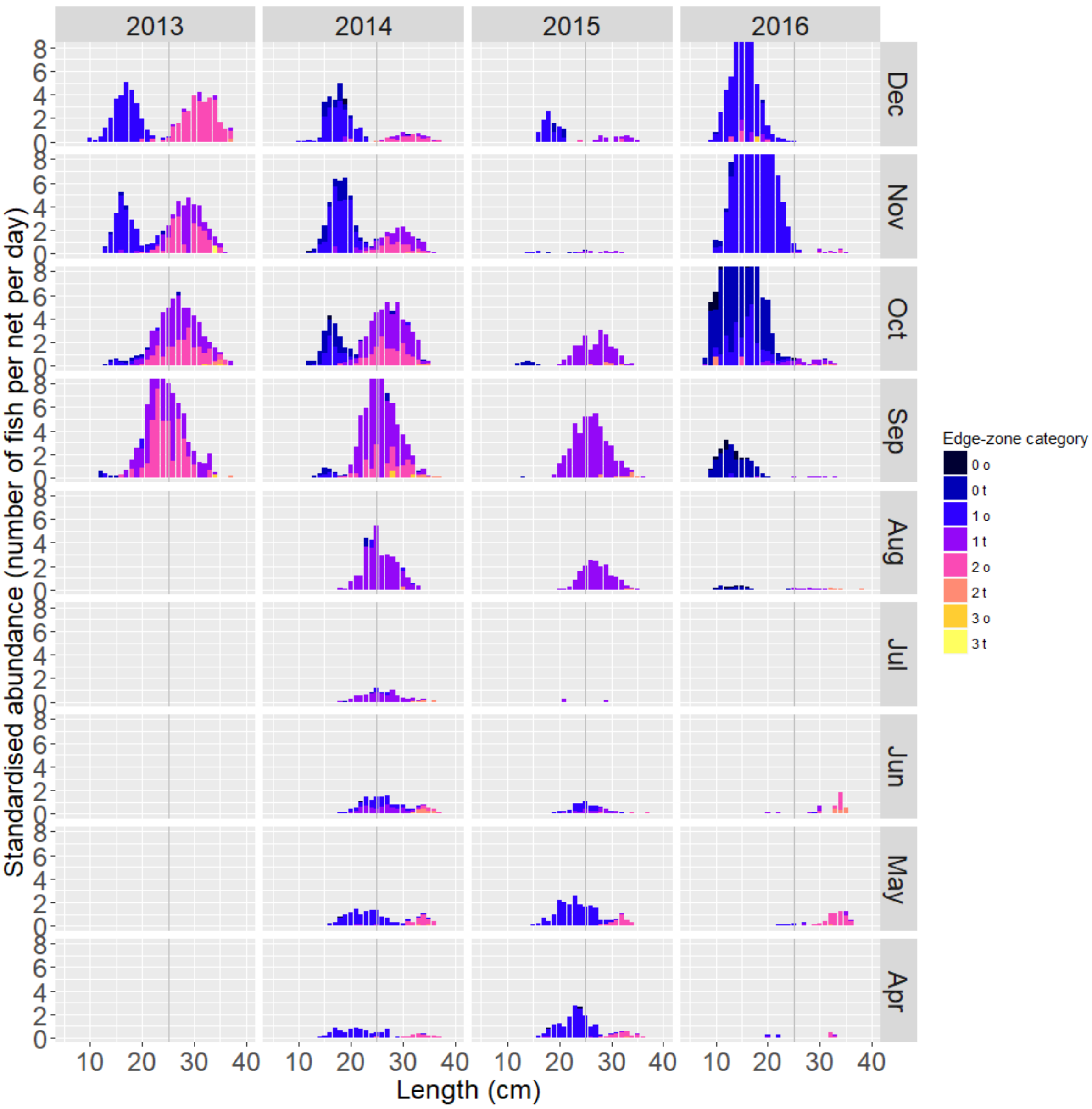


Figure 6: Proportions of edge-zone category per length class overlaid ono the standardised length frequencies

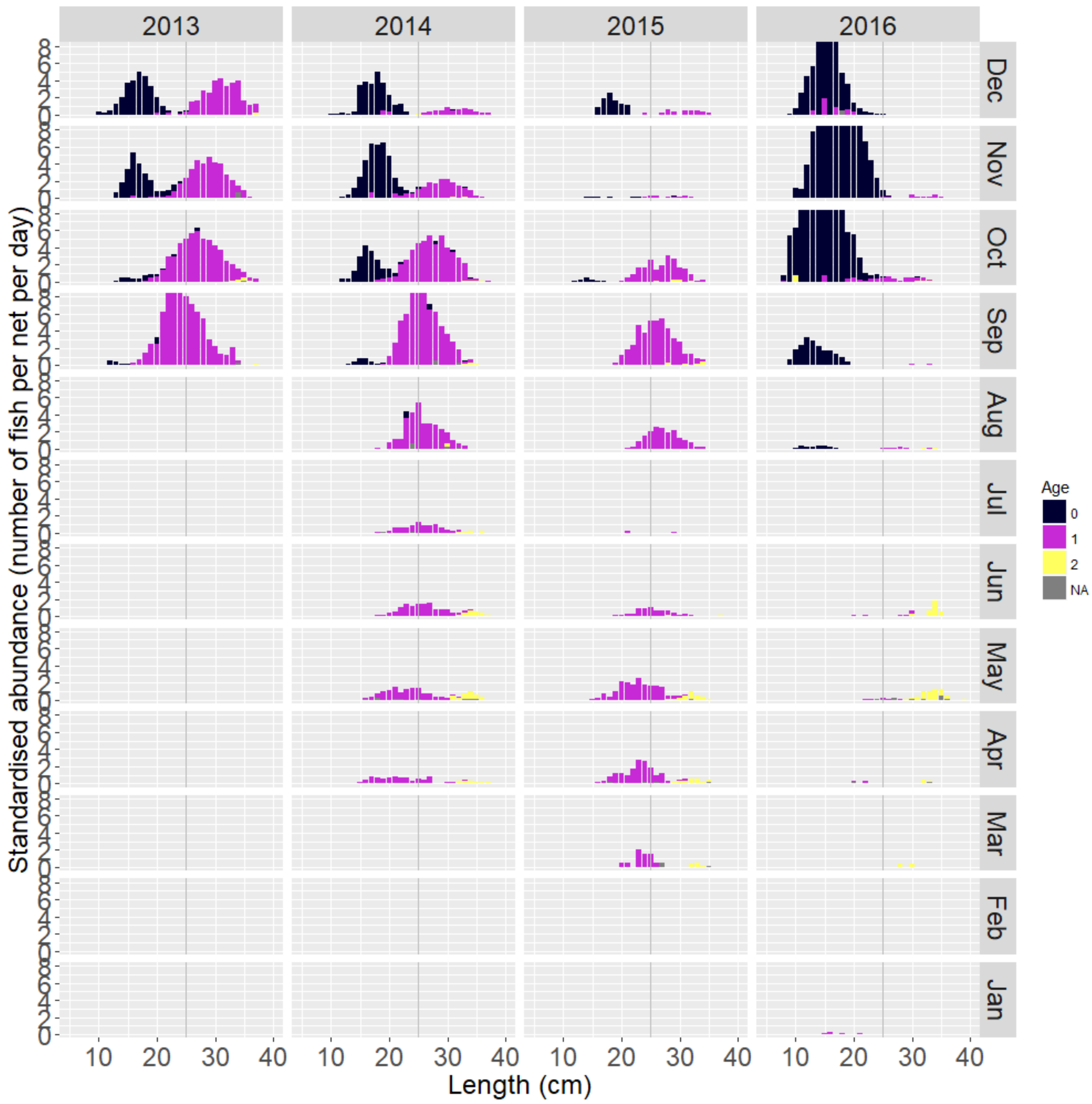


Figure 7: Proportion of ages per length class assigned following the conditions in Table 1, overlaid onto the standardised length frequencies. Otolith analysis confirms that the smaller cohort visible in the samples between September and December is age-0, and the larger cohort is age-1.

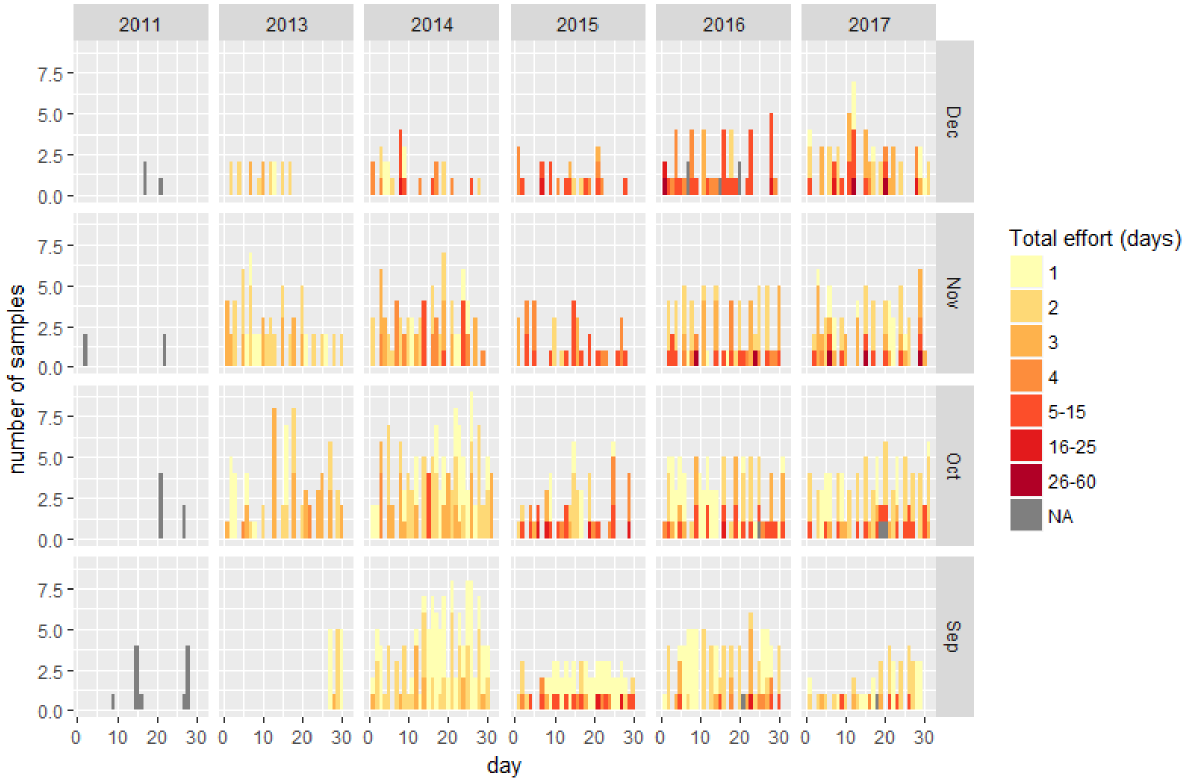


Figure 8: The distribution of sampling effort across years and months. The coloured bars indicate the number of samples which were collected each day of each month each year. The total effort key indicates the number of days soaking multiplied by the number of nets which contributed to the sample, as it was often the case that different sampling events did not represent equal fishing effort.

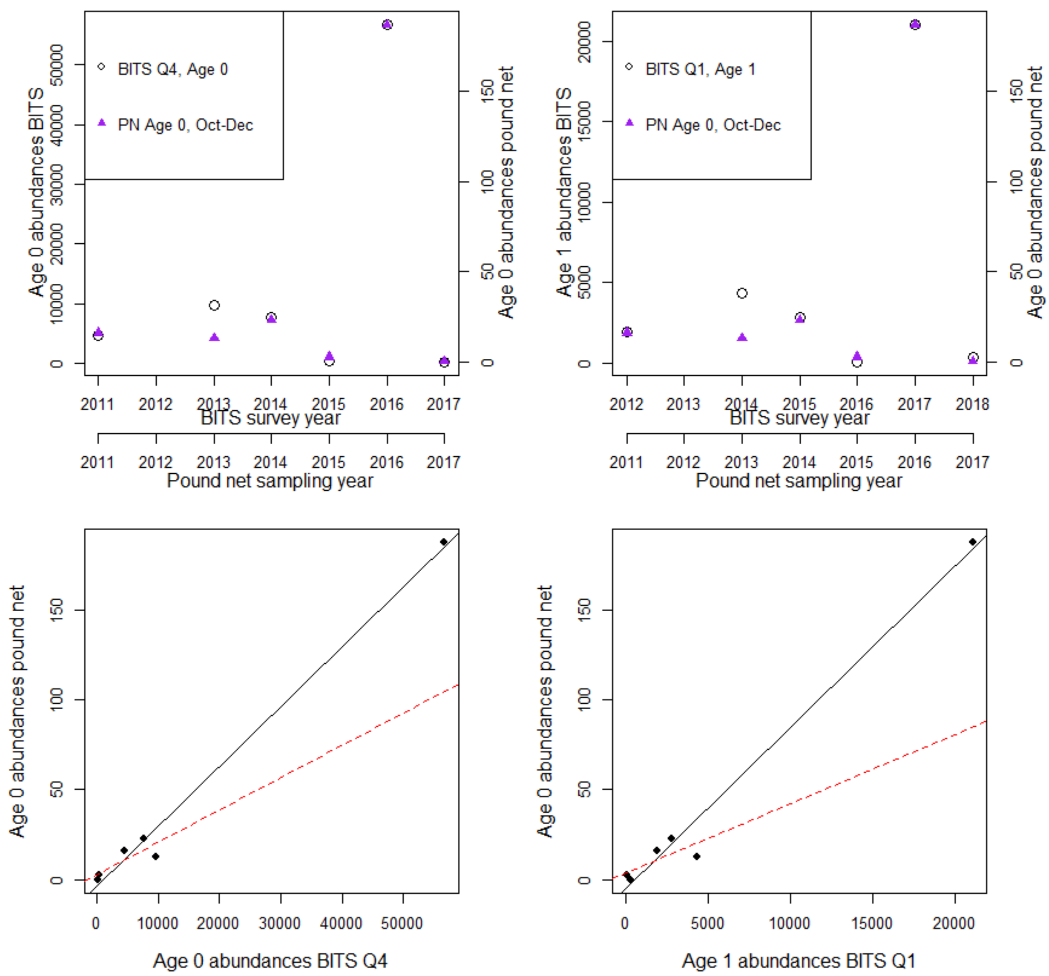


Figure 9: The relationship between average daily abundances of age-0 WBC in the pound nets (PN) during October-December, and the abundances of age-0 WBC estimated from BITS in Q4 (left plots) or the abundances of age-1 WBC from BITS in Q1 (right plots). These models produced the highest R^2 values of all data combinations tested (Table 3). The black regression line in the bottom row of plots indicates the relationship between the two datasets, including data from 2016. The red dashed regression line illustrates the relationship if 2016 is excluded from analysis (Table 3).