INTER-BENCHMARK PROCESS ON BALTIC SPRAT (*SPRATTUS SPRATTUS*) AND HERRING (*CLUPEA HARENGUS*) (IBPBASH)

VOLUME 2 | ISSUE 34

ICES SCIENTIFIC REPORTS

RAPPORTS SCIENTIFIQUES DU CIEM

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

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

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

ISSN number: 2618-1371 I © 2020 International Council for the Exploration of the Sea

ICES Scientific Reports

Volume 2 | Issue 34

INTER-BENCHMARK PROCESS ON BALTIC SPRAT (*SPRATTUS SPRATTUS*) AND HERRING (*CLUPEA HARENGUS*) (IBPBASH)

Recommended format for purpose of citation:

ICES. 2020. Inter-Benchmark Process on BAltic Sprat (*Sprattus sprattus*) and Herring (*Clupea harengus*) (IBPBash).
ICES Scientific Reports. 2:34. 44 pp. http://doi.org/10.17895/ices.pub.5971

Editors

Bjarte Bogstad

Authors

Mikaela Bergenius • Margit Eero • Simon Fischer • Tomas Gröhsler • Jan Horbowy • Marc Taylor

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

Contents

i Executive summary

The Inter-Benchmark Process (IBP) on BAltic Sprat (*Sprattus sprattus*) and Herring (*Clupea harengus*) (IBPBASH) was carried out by three WebEx meetings in March 2020. The process had two purposes;

1. to evaluate the appropriateness of the use of the natural mortality estimates derived from the most recent (2019) multispecies SMS keyrun for the Baltic in the stock assessments for herring and sprat;
2. To re-examine and update MSY and PA reference points for both stocks according to ICES guidelines

It was found appropriate to use unsmoothed values of natural mortality from the SMS keyrun in the assessments for both stocks. Future M values should be predicted using a model which includes cod spawning stock biomass.

Both MSY and PA reference points were re-estimated for both stocks. For sprat, the biomass reference points were unchanged, while all fishing mortality reference points increased. For herring, the biomass reference points were lowered by about 25%. F_{MSY} and the corresponding range were practically unchanged, while F_{lim} and F_{pa} increased slightly. The regime shift in the Baltic around 1990 is a challenge when deciding which part of the time-series to be used, this applies to both stocks.

ii Expert group information

Expert group name	An Inter-Benchmark Process (IBP) on BAltic Sprat (*Sprattus sprattus*) and Herring (*Clupea harengus*) (IBPBASH)
Expert group cycle	Annual
Year cycle started	2020
Reporting year in cycle	1/1
Chair	Bjarte Bogstad, Norway
Meeting venue and dates	WebEx meetings 6 ,16 and 24 March 2020, ten participants

1 Introduction

Terms of reference

An Inter-Benchmark Process (IBP) on BAltic Sprat (*Sprattus sprattus*) and Herring (*Clupea harengus*) (IBPBASH), chaired by Bjarte Bogstad, Norway, and attended by two invited external experts Simon Fischer, UK and Marc Taylor, Germany will be established and will work by correspondence on 6th and 16th March 2020 to:

a) Evaluate the appropriateness of the use of the natural Mortality estimates derived from the multispecies SMS keyrun for the Baltic (WGSAM 2019) in the stock assessments for herring and sprat;
b) Update the stock annex as appropriate;
c) Re-examine and update MSY and PA reference points according to ICES guidelines (see Technical document on reference points);

Stocks	Stock leader
Sprat (*Sprattus sprattus*) in Subdivisions 22–32 (Baltic Sea)	Jan Horbowy
Herring (*Clupea harengus*) in subdivisions 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea)	Tomas Gröhsler

The IBP will report by 3rd April 2020 for the attention of ACOM.

The meeting was conducted by WebEx, with meetings 6, 16 and 24 March 2020.

ToR a) is addressed in Section 2 and 3 and ToR c) in Sections 4 and 5, further updated stock annexes are given as appendices to the report.

2 Natural mortality in assessment

2.1 Sprat

Comparison of predation mortality, M2

The sprat stock assessment, prediction of recruitment, and prediction of biomass and catch development conducted during WGBFAS in 2019 (ICES, 2019a) were repeated with the same data and parameterisation as at WGBFAS, except natural mortality. New natural mortality estimates obtained from new SMS run (ICES, 2019b) were used in the analysis and results of new assessment and predictions were compared with results obtained at WGBFAS in 2019.

Figure 2.1a presents comparison of new estimates of predation mortality with estimates used in previous assessment. Average values of M2 (age 1 –7) are similar to the values used previously. In Figure 2.1b, previous and new M2 are compared in relative terms by age. Relative differences between both series fluctuate usually within +/-30% for ages 2 and older. Bigger differences are between M2 estimates for ages 0–1; new estimates are higher than old ones by on average 60% for age 0 and 45% for age 1. Age 0 is not used in the assessment.

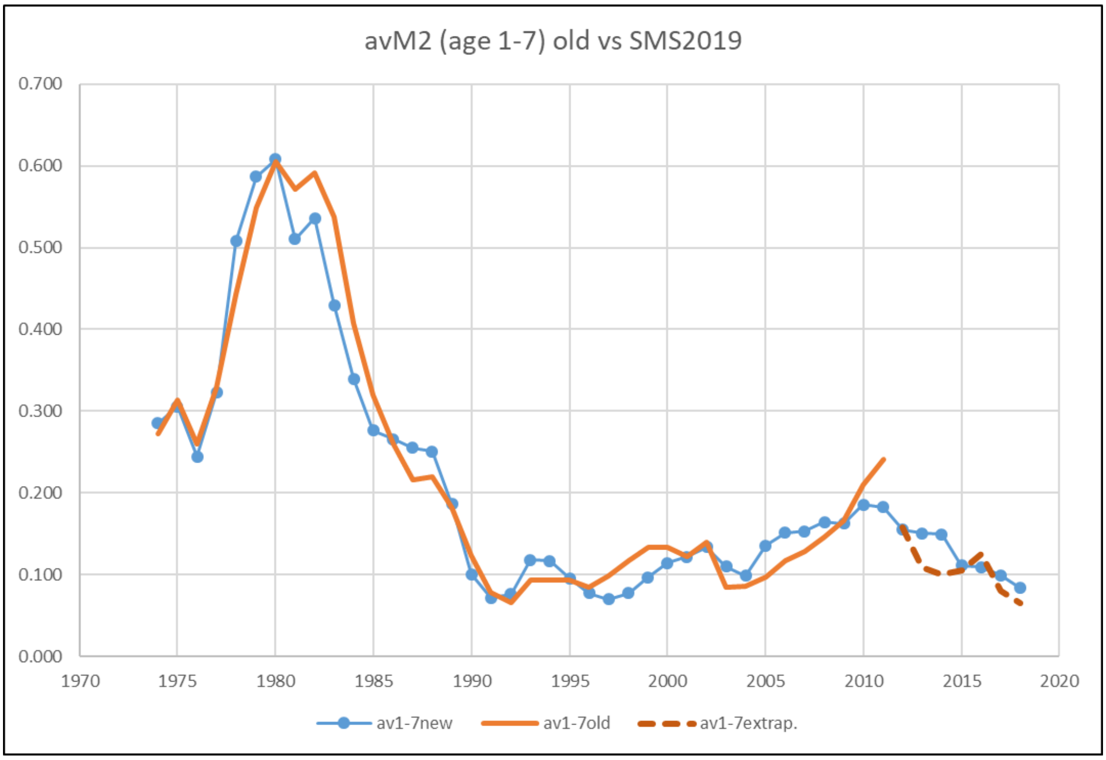

Figure 2.1a. Average (age 1–7) M2 from old & new SMS; in addition extrapolations of M2 for 2012 onwards.

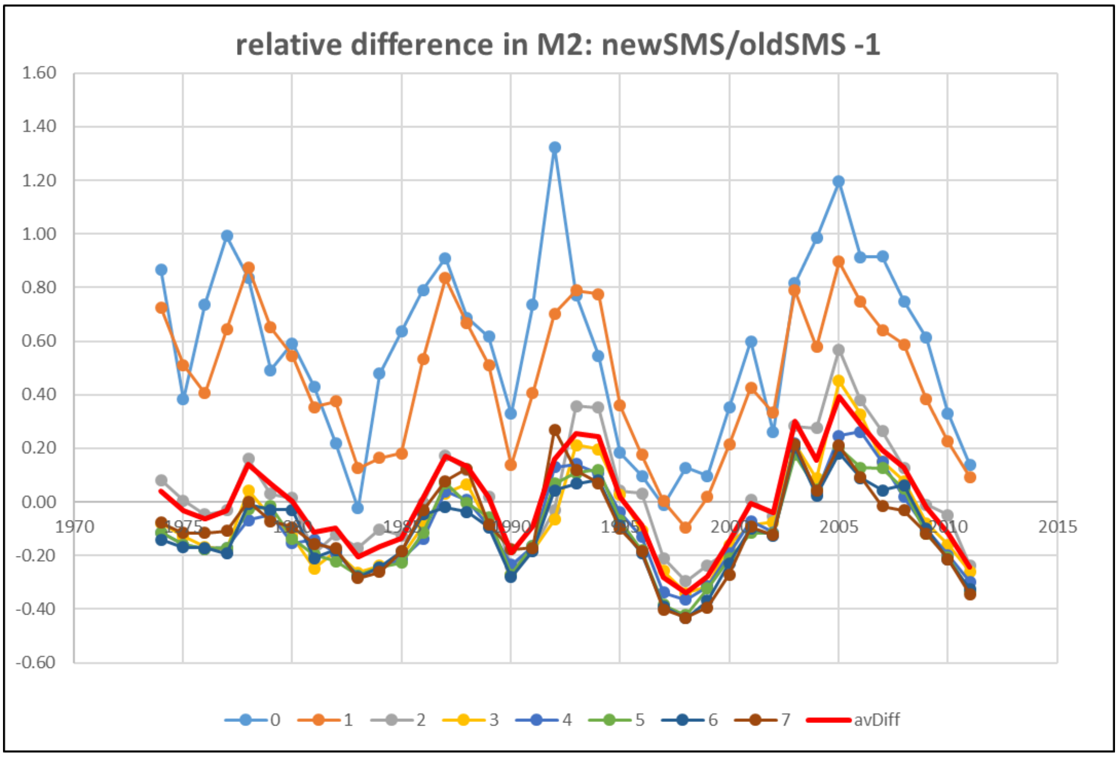

Figure 2.1b. Relative difference at age (0–7) between M2 from old & new SMS.

Comparison of assessments

The estimates of stock size, fishing mortality, and recruitment with new M values are very similar to last year estimates of the WGBFAS (Figure 2.2). Most of differences did not exceed a few percent; somewhat higher differences were observed at the beginning of time-series (up to early 1980s) for SSB and F. Diagnostics of the XSA with new M (in terms of mean log catchabilities, their standard errors, correlations between survey and XSA abundances, survivors estimates, and their weights) was also very similar to diagnostics from last WGBFAS. In Figure 2.3 mean log catchabilities and their standard errors from both runs are shown.

Retrospective estimates of SSB, F_{bar}, and recruitment are also similar to obtained in previous assessment (Figure 2.4). Mohn's rho for SSB, F_{bar}, and recruitment were 0.17, -0.2, and 0.04, respectively.

The XSA model provides standard errors of survivors' estimates, but they include only variability in survey indices, the catches in XSA are treated as exact values. To consider uncertainty in the catches parametric bootstrap was performed and confidence intervals for biomass estimates were obtained. Catches were distorted by log-normal random error with standard deviation of 0.2 and 0.3. Errors within years were assumed correlated. The median biomass and its 90% confidence intervals for log catch sd of 0.3 are presented in Figure 2.5. The variability of estimates in terminal year are equivalent to CV=0.15.

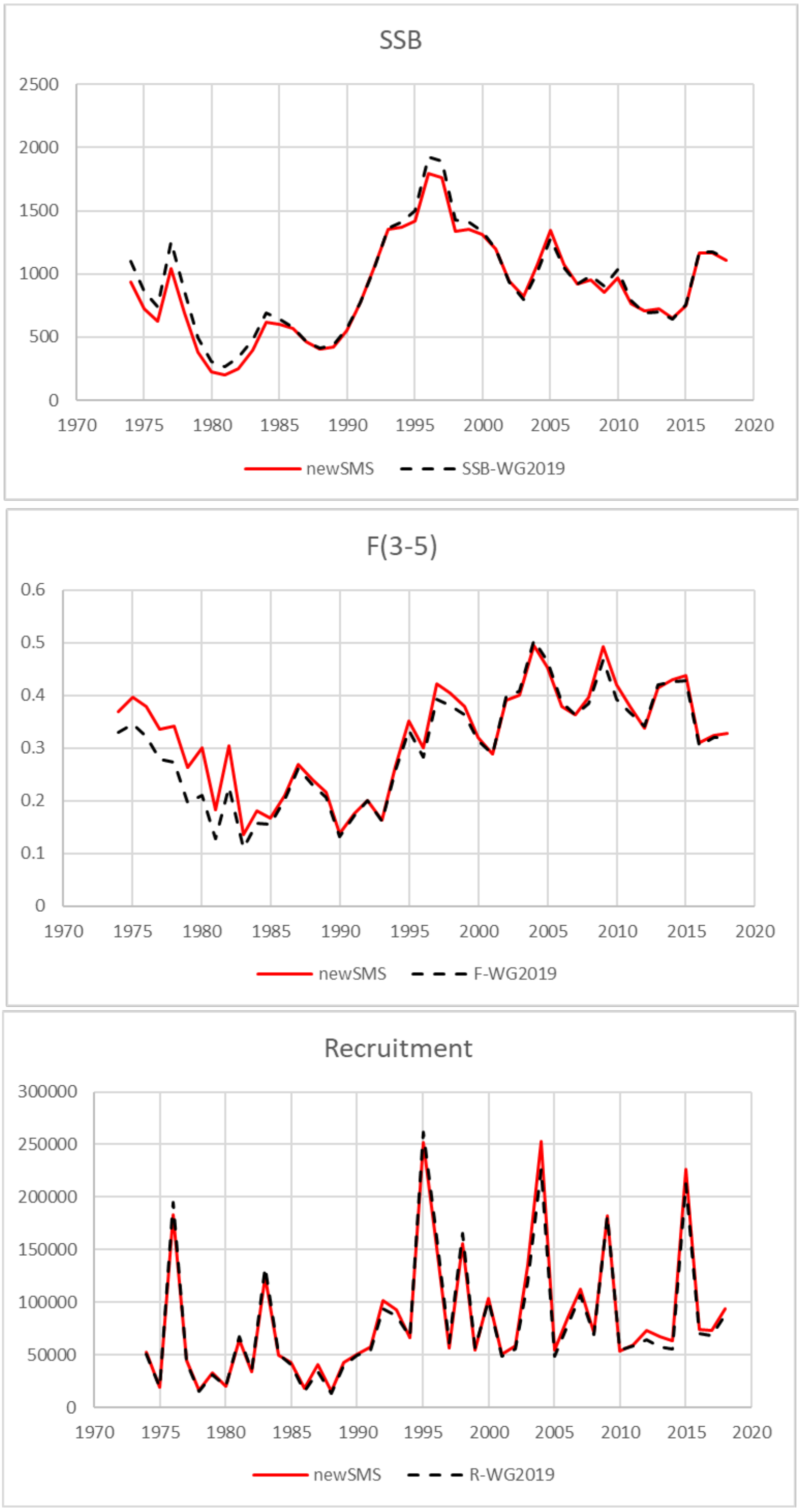

Figure 2.2. Estimates of SSB, fishing mortality, and recruitment from assessment with new M and assessment at WGBFAS 2019 (ICES, 2019a).

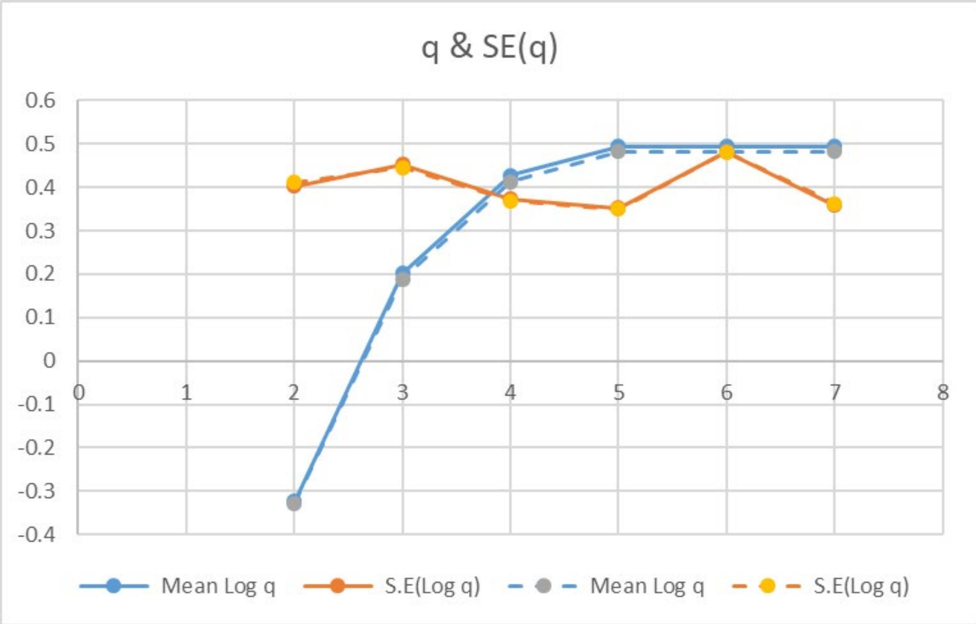

Figure 2.3. Diagnostics for new (solid line) and old (broken line) assessment exemplified as mean log catchabilities and their standard errors.

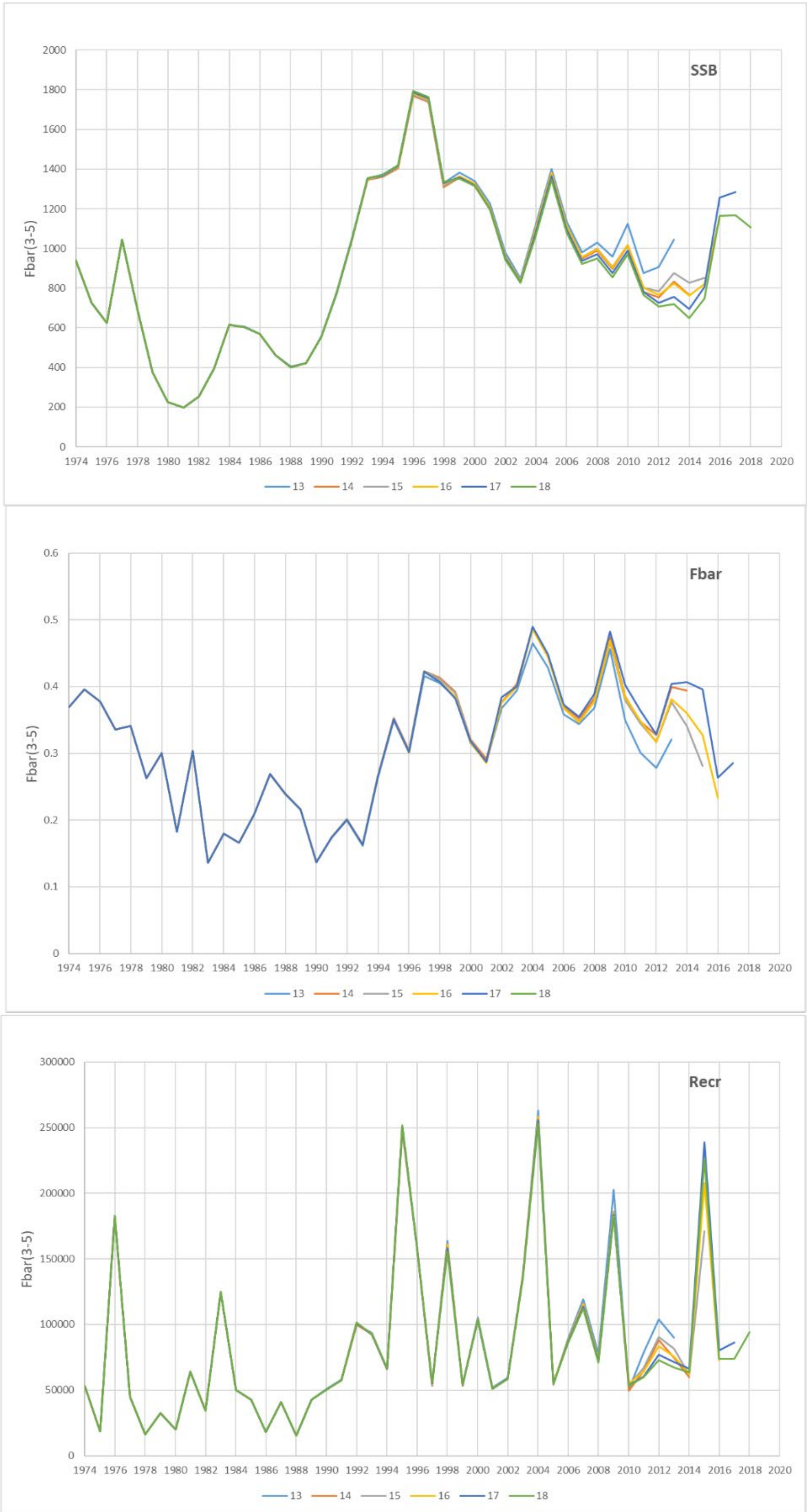

Figure 2.4. Retrospective estimates of SSB, F_{bar}, and recruitment in assessment using new M values.

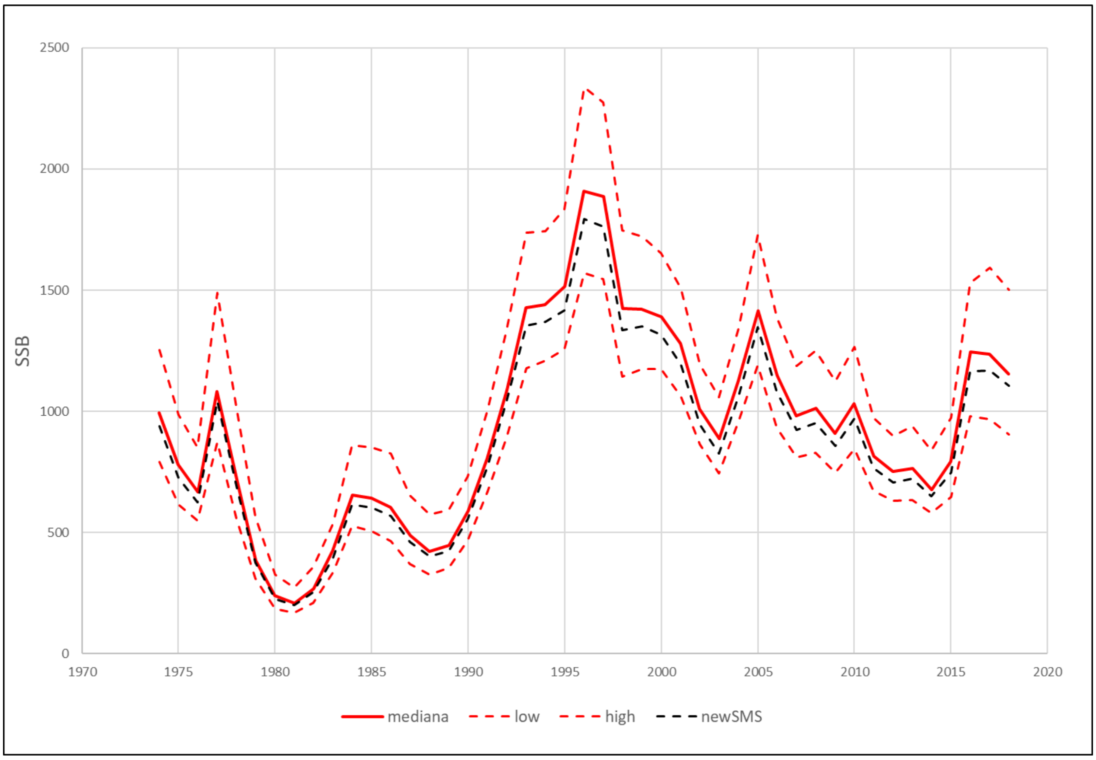

Figure 2.5. The median SSB and bootstrap estimates of 5 and 95 percentiles of its distribution (red). Biomass estimated in deterministic run is given for comparison (black).

Comparison of recruitment predictions

The RCT3 software has been to predict recruitment-at-age 1 in intermediate year; that estimate is next used in the short-term predictions. As could be expected from similarities in recruitment estimates with old and new M, the RCT3 estimates using data from new assessment are very similar to values estimated using old M data (Figure 2.6).

As both assessments and RCT3 estimates in present assessment are very similar to estimates obtained by WGBFAS in 2019, the short-term projections were expected to be similar also (Figure 2.7). Differences between projected catches and biomasses were at a level of 2%.

Conclusions

Assessment and predictions using M from new SMS are very similar to assessment and predictions performed at WGBFAS in 2019. The later used M from previous SMS till 2012 and M extrapolated basing on relationship between cod SSB and M in next years.

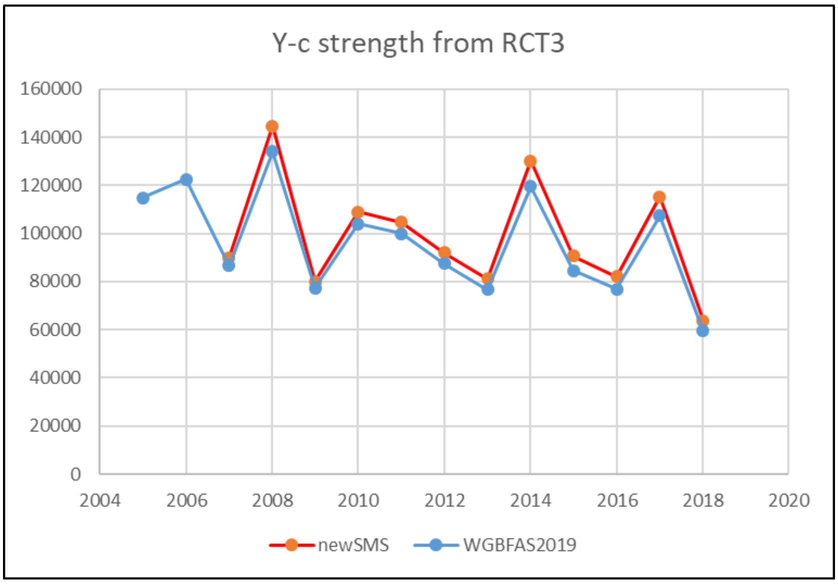

Figure 2.7. Recruitment (age 1) estimates at intermediate year using RCT3 software in present assessment and obtained by WGBFAS in 2019.

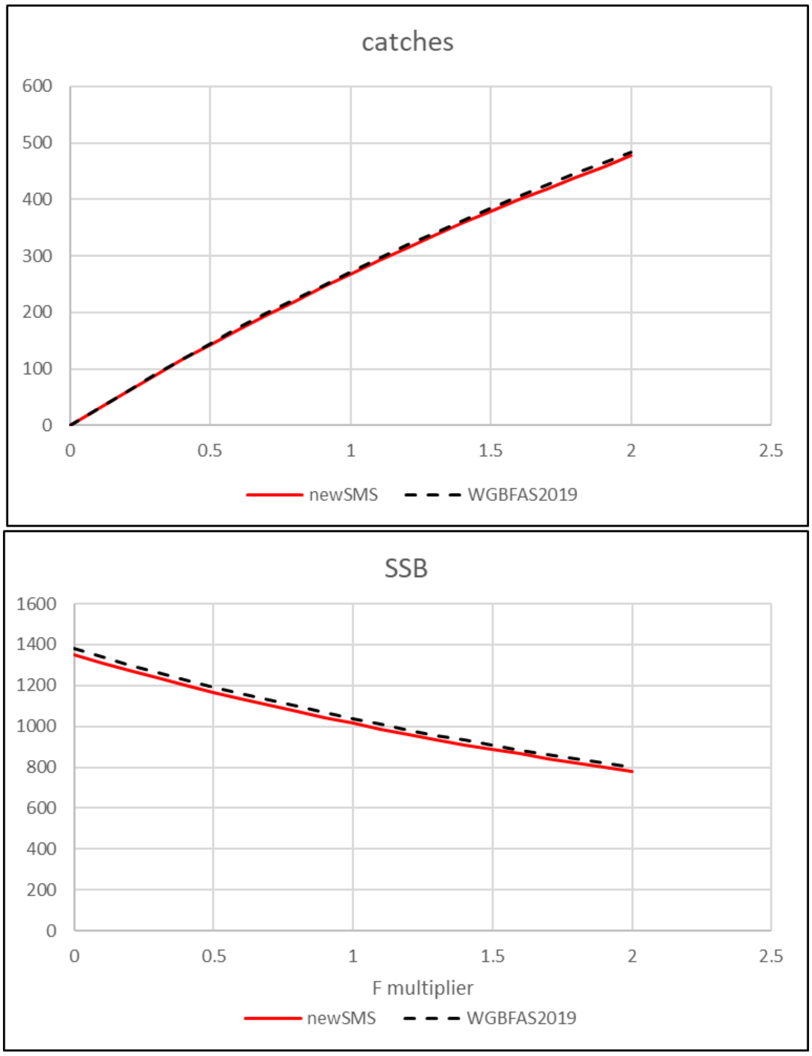

Figure 2.8. Comparison of short-term predictions performed in present analysis with new M2 data and obtained by WGBFAS in 2019.

2.2 Herring

Comparison of predation mortality, M2

The central Baltic herring (CBH) stock assessment conducted during WGBFAS in 2019 (ICES, 2019a) were repeated with the same data and parameterisation as at WGBFAS, except natural mortality. New natural mortality estimates obtained from new SMS run were used in the analysis and results of new assessment were compared with results obtained at WGBFAS in 2019. Further details are given in WD by Gröhsler and Neuenfeldt.

IBPBASH decided to use unsmoothed M values from WGSAM (ICES, 2019b) for the entire period following the same procedure as during the last update of M values in 2012 (WGSAM, ICES, 2012). The comparison of new estimates of predation mortality with estimates used in previous assessment (Figures 2.9 and 2.10) showed the following overall differences:

- The new estimates of age 1 are in most years far higher and positive.
- In general the new M values of age 2–8+ give higher and positive estimates at beginning of the time-series till around the mid-1980s where they all become more or less negative.
- The new estimates of M show a decreasing trend in last years.

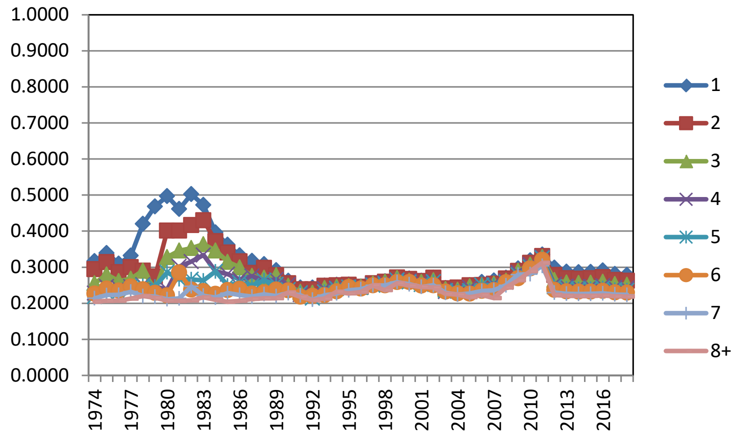

Figure 2.9a. M-at-age as used by WGBFAS in 2019.

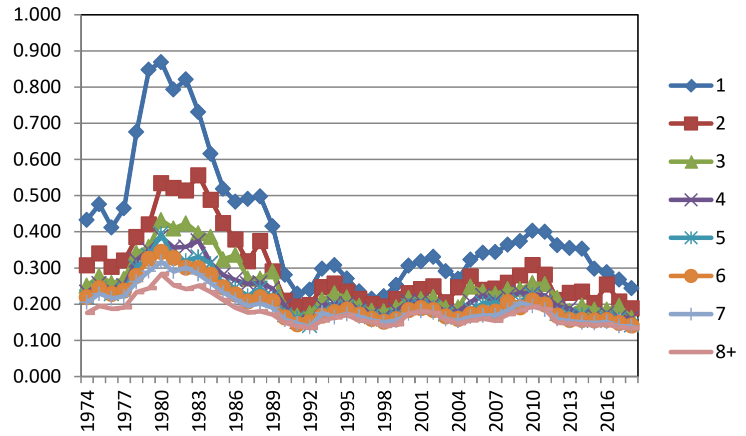

Figure 2.9b. New M-at-age as derived from SMS (WGSAM 2019).

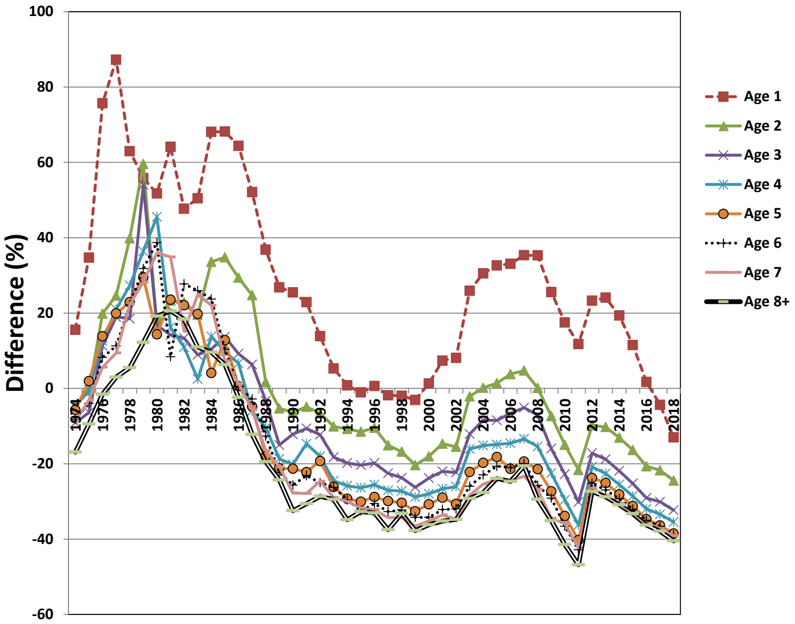

Figure 2.10. Difference (%) of new M-at-age derived from SMS (WGSAM 2019) compared to M used by WGBFAS 2019.

Comparison of assessments

The estimates of stock size (Figure 2.11), fishing mortality (Figure 2.12), and recruitment (Figure 2.13) with new M values are largely consistent with previous stock trends, but deviate increasingly since the early 1980s compared to WGBFAS 2019. The new M values gave a more optimistic part of the stock development at the beginning the time-series (1974–1982: higher SSB/Recruitment and lower F), whereas the stock development since 1984 show a more pessimistic view of the stock status (lower SSB/Recruitment and higher F), The differences (%) in SSB compared to the final run of WGBFAS in 2019 vary between >+30% at the beginning, no difference in 1984 and >-20 % at the end of the time-series.

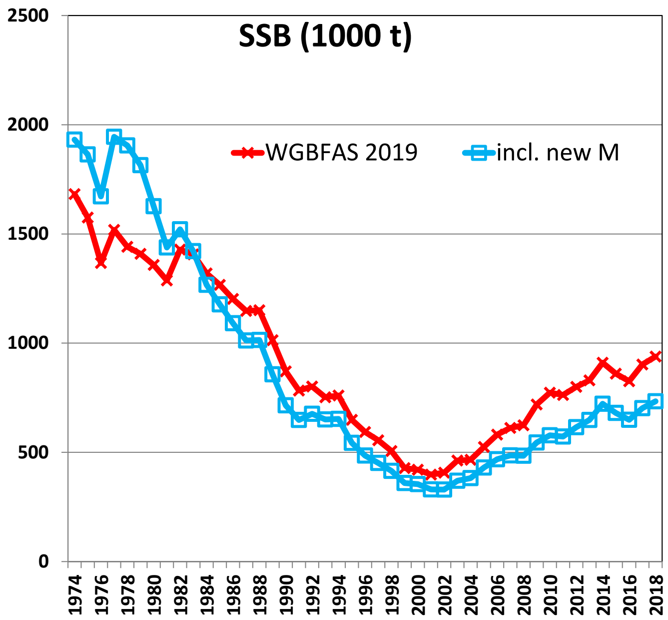

Figure 2.11a. Comparison of SSB (1000 t) of the:

- Final run of WGBFAS 2019,
- Run incl. new M (WGSAM 2019),

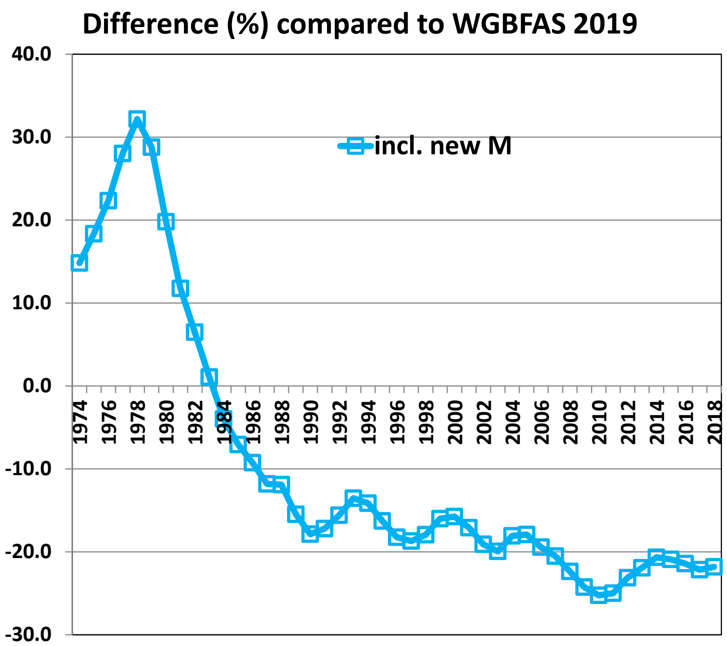

Figure 2.11b. Corresponding differences (%) in SSB compared to the final run of WGBFAS 2019.

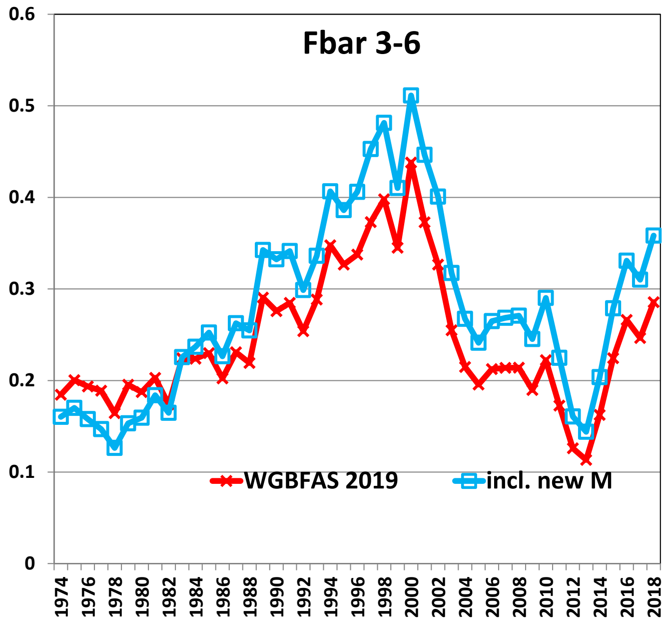

Figure 2.12a. Comparison of F_{bar} (3–6) of the:

- Final run of WGBFAS 2019,
- Run incl. new M (WGSAM 2019).

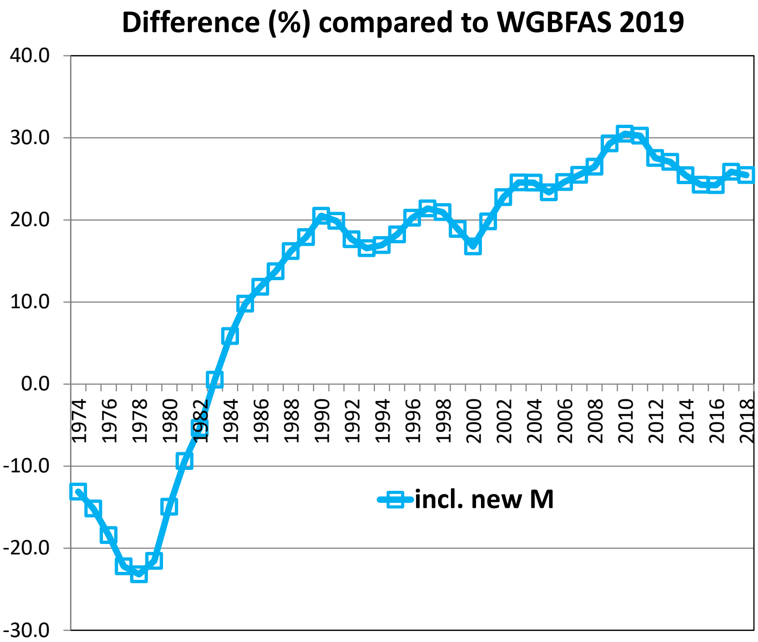

Figure 2.12b. Corresponding differences (%) in F_{bar} (3–6) compared to the final run of WGBFAS 2019.

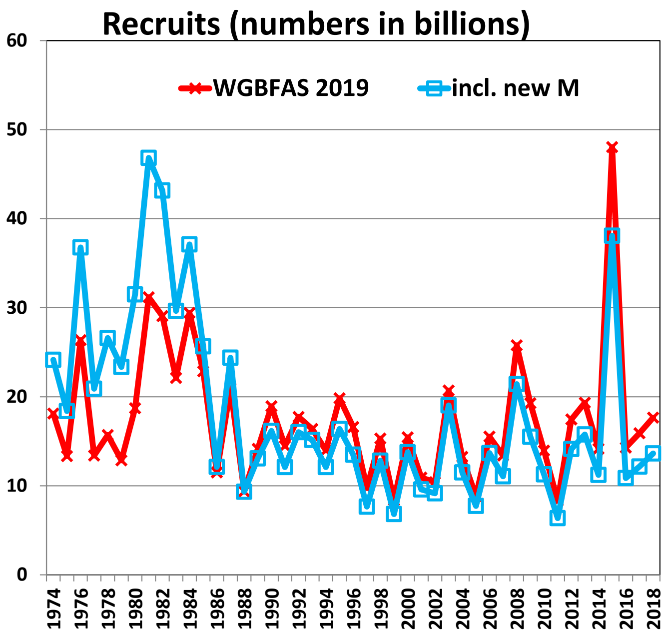

Figure 2.13a. Comparison of Recruitment-at-age 1 of the:

- Final run of WGBFAS 2019,
- Run incl. new M (WGSAM 2019),

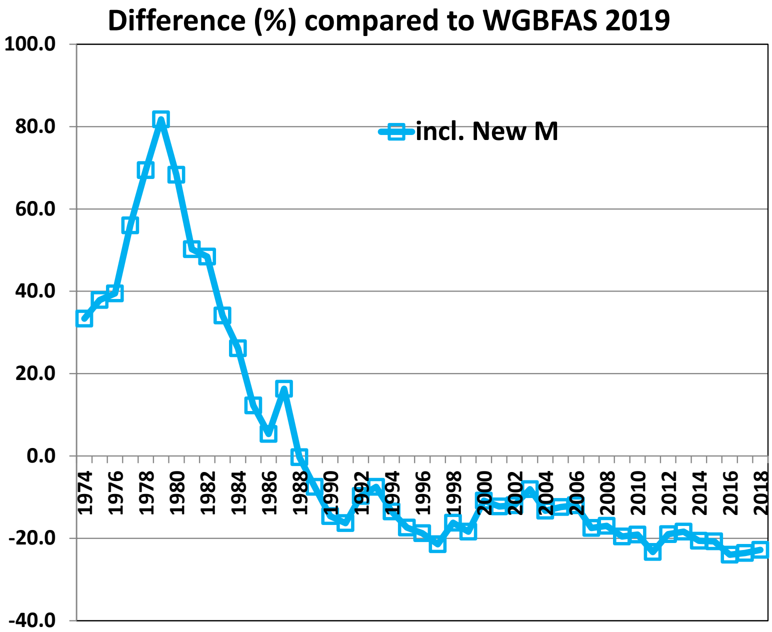

Figure 2.13b. Corresponding differences (%) in Recruitment-at-age 1 compared to the final run of WGBFAS 2019.

Diagnostics

Diagnostics of the XSA with new M (in S.E. (log q) and further regression statistics) was similar to diagnostics from last WGBFAS. Figure 2.14 shows the S.E. (log q) at-ages 2–7 of both runs.

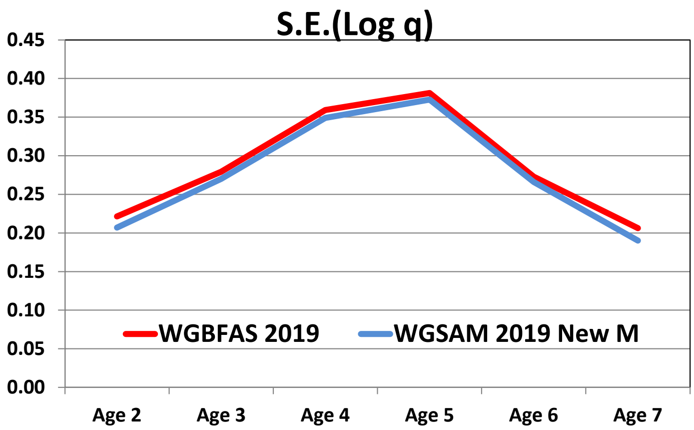

Figure 2.14a. Comparison of S.E. (log q) at-ages 2–7 of the:

- Final run of WGBFAS 2019,
- Run incl. new M.

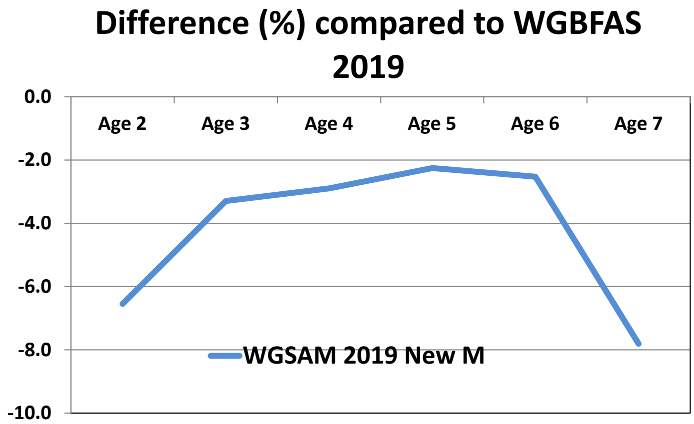

Figure 2.14b. Corresponding differences (%) of S.E (log q) compared to the final run of WGBFAS in 2019.

Baltic International Acoustic Survey (BIAS): Log catchability residuals

Figure 2.15 shows the log-catchability residuals at-age of the Baltic International Acoustic Survey (BIAS) of the a) Final run of WGBFAS in 2019 and b) Run incl. new M.

Compared to the final run of WGBFAS in 2019 the runs with new M values at-age resulted in equal or even slightly lower residuals.

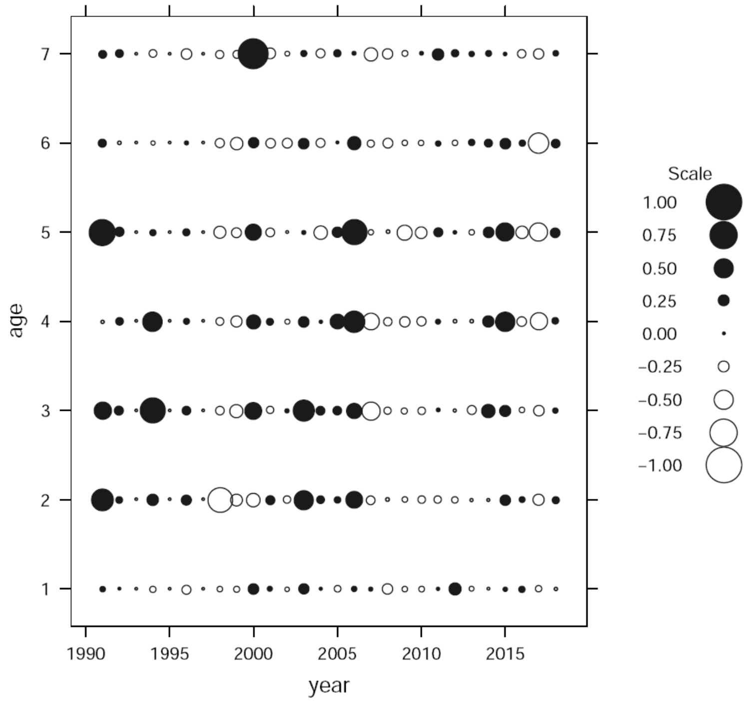

Figure 2.15a. Baltic International Acoustic Survey (BIAS): Log catchability residuals at age of the Final run of WGBFAS 2019.

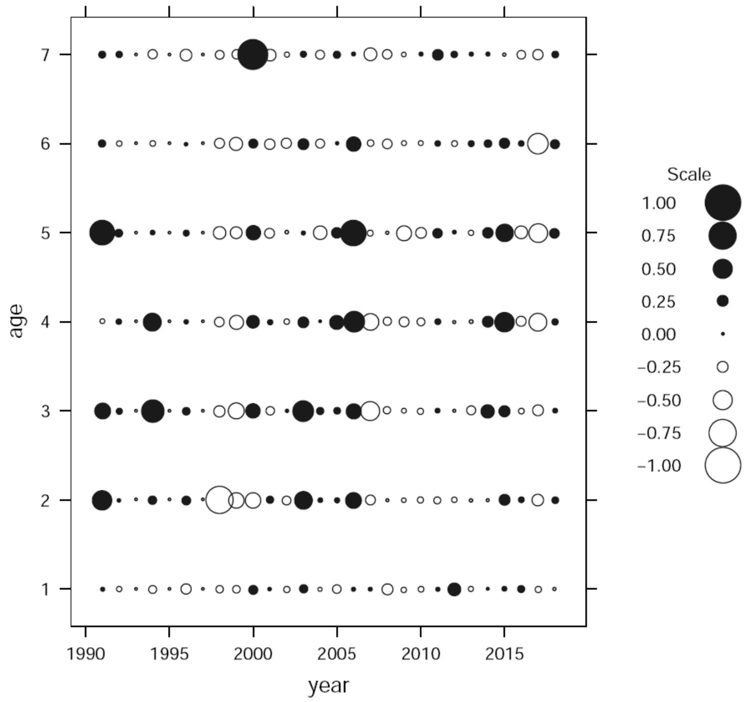

Figure 2.15b. Baltic International Acoustic Survey (BIAS): Log catchability residuals at-age of the run incl. new M.

Retrospective Analysis

Retrospective estimates of SSB, F_{bar}, and recruitment are also similar to the values obtained in previous assessment (Figure 2.16). Mohn's rho for SSB, F_{bar}, and recruitment were:

Mohn's rho	SSB	F_{bar}	Recruitment
WGBFAS 2019	0.06681156	-0.04920581	-0.06916053
New M	0.07885813	-0.05538882	-0.03824129

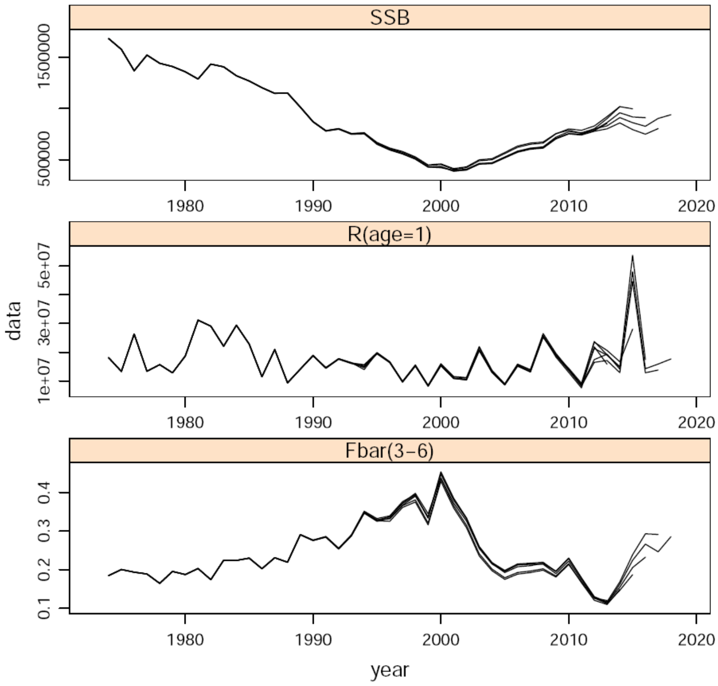

Figure 2.16a. Retrospective analysis of the final run of WGBFAS 2019.

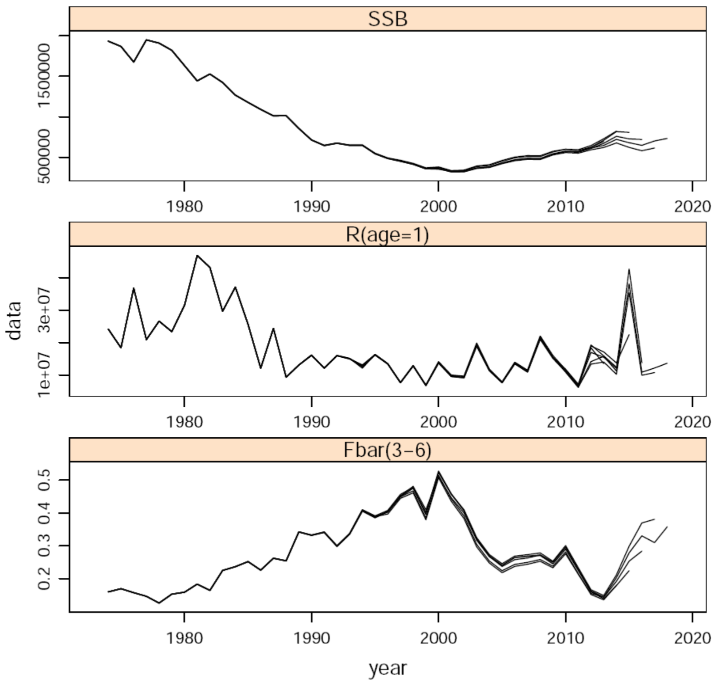

Figure 2.16b. Retrospective analysis of the run incl. new M (WGSAM 2019).

Conclusion/summary

The run based on the new M values resulted in:

- a more optimistic part of the stock development at the beginning the time-series (1974–1982: higher SSB/Recruitment and lower F), whereas the stock development since 1984 show a more pessimistic view of the stock status (lower SSB/Recruitment and higher F);
- equal or slightly improved diagnostics;
- equal or even slightly lower log catchability residuals;
- rather similar retrospective pattern.

3 Natural mortality-updating in future

3.1 Sprat

SMS is expected to be updated every few years, so there is a need to predict natural mortality (M=M1+M2, where M1 is residual natural mortality and M2 is predation mortality) in years when new estimates from SMS are not available. One possibility is to predict M from cod biomass. Two potential predictors of M are available, SSB and biomass of cod >35 cm, both routinely now provided by WGBFAS (ICES, 2019a). Figure 3.1 show regression of average M from SMS against cod biomass for years 1974–2018. SSB is somewhat better predictor of M (R^2=0.9) than biomass of cod>35 cm (R^2=0.8). When both regressions are constrained to a shorter time period, i.e. years from 1990 onwards, the predictive power of cod biomass markedly declines, being zero for biomass of cod >35 cm. It may be noted that when SSB is taken as a predictor, the regression parameters for shorter period are not very different from parameters of regression for whole period.

Other option of predicting M in years for which SMS was not updated could be assuming M at average level of recent years, but that was considered less reliable when cod biomass undergoes changes.

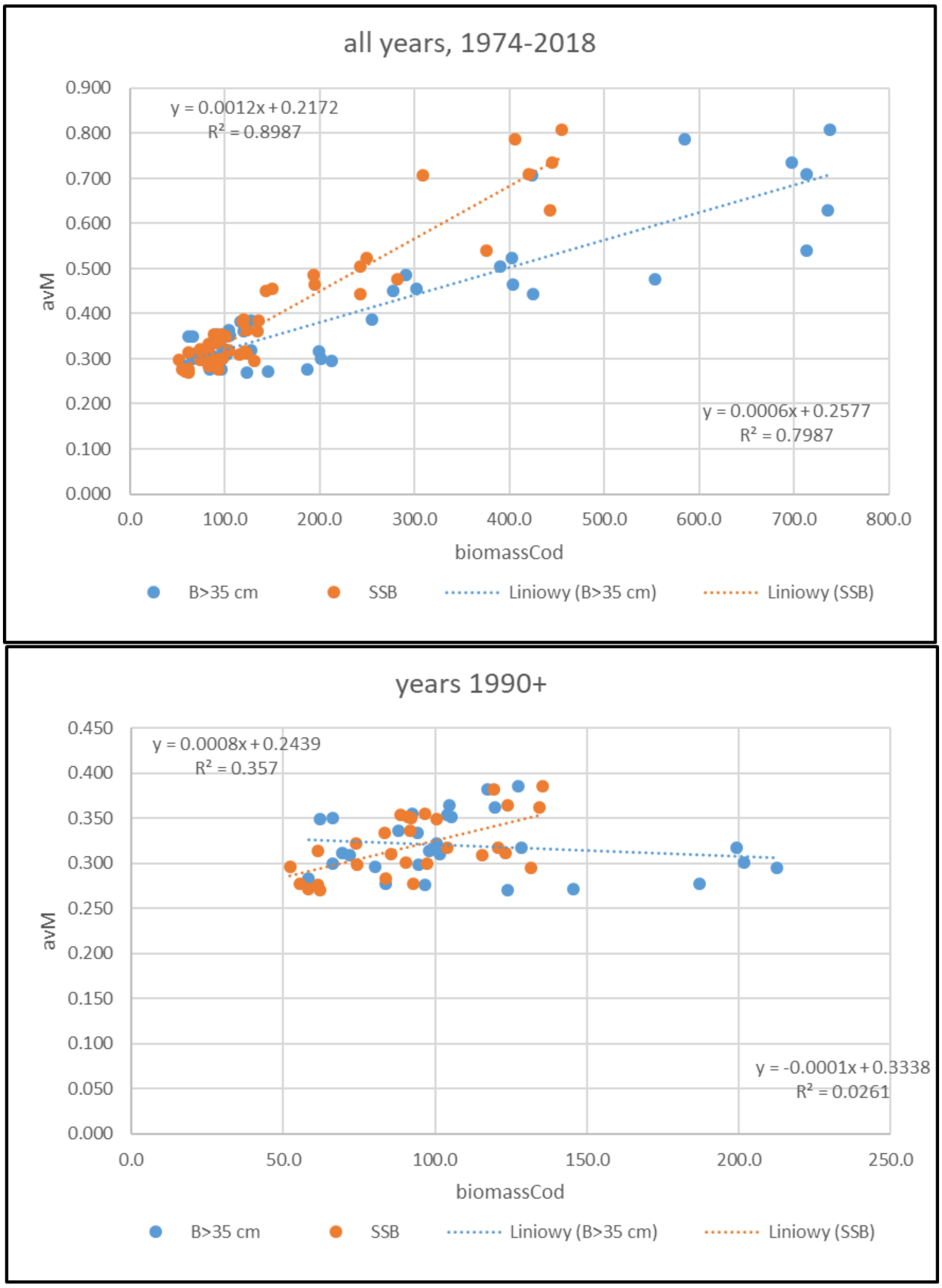

Figure 3.1. The dependence of average natural mortality, avM, on cod biomass (SSB and biomass of cod>35 cm) for all data years and for data from 1990 onwards.

3.2 Herring

The prediction of M for herring in years when new estimates from SMS are not available may be done in similar way as suggested in case of sprat. Figure 3.2 shows regression of average M (M=M1+M2, M1=0.1) from SMS against cod biomass for years 1974–2018. SSB is a slightly better predictor of M2 (R^2=0.93) than biomass of cod>35 cm (R^2=0.88). When both regressions are constrained to data for years from 1990 onwards, the predictive power of cod biomass markedly declines, being zero for biomass of cod >35 cm. It may be noted that when SSB is taken as predictor, the regression parameters for shorter period are almost identical to parameters of regression for the whole period.

Other option of predicting M in years for which SMS was not updated could be assuming M at average level of recent years, but that was considered less reliable when cod biomass undergoes changes.

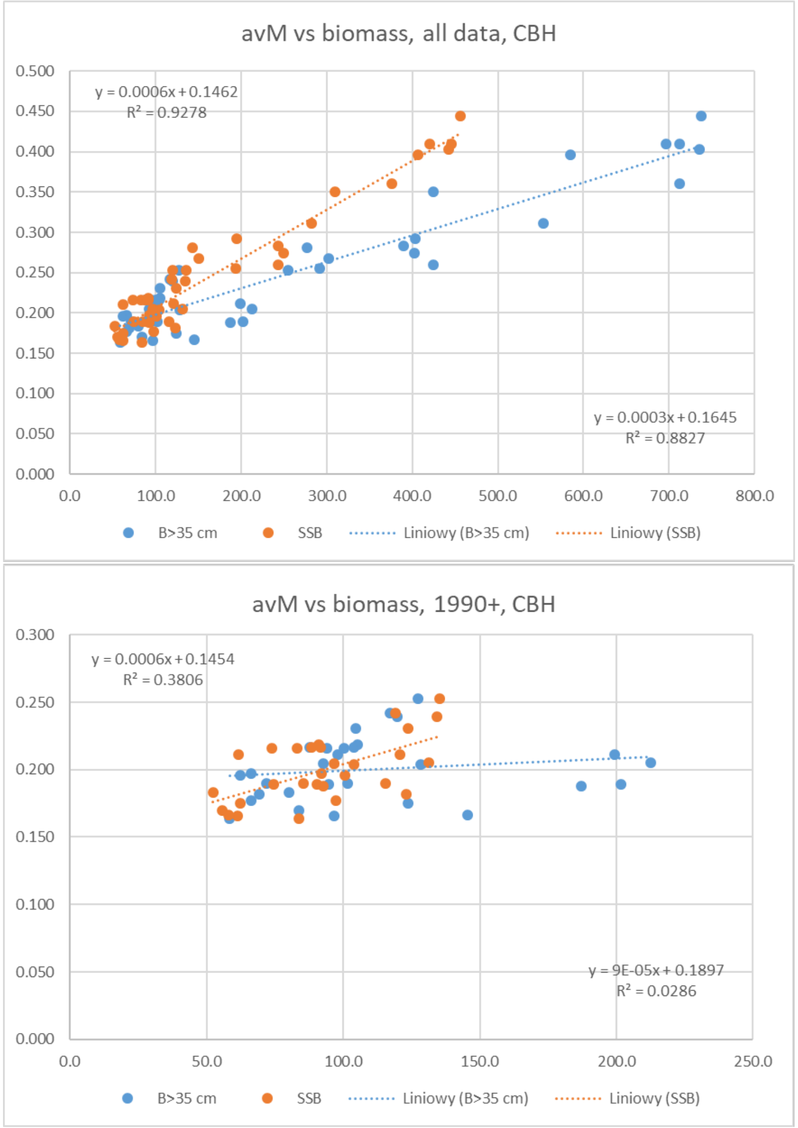

Figure 3.2. The dependence of average natural mortality, avM, on cod biomass (SSB and biomass of cod>35 cm) for all data years and for data from 1990 onwards.

4 Sprat reference points

4.1 B_{lim} and B_{pa}

As in former approaches (benchmark workshop WKBALT, ICES (2013)) the B_{lim} was estimated as the biomass which produces 50% of maximal recruitment from S–R model following the Myers *et al*. (1994) approach. Fits of both Beverton and Holt and Ricker S–R relationships were similar in terms of AIC, which was -7.62 and -6.68 for Beverton and Holt and Ricker S–R model, respectively (Figure 4.1). The segmented regression fit was worse with AIC=-5.08. The biomass which produces 50% of maximum recruitment in Beverton and Holt model was 470 kt, and 345 kt for Ricker S–R model. The average of both values is 407.5 kt which was rounded to 410 kt and accepted as B_{lim} value. That estimate is identical to B_{lim} used up to now and estimated at previous benchmark (ICES, 2013).

B_{pa} was estimated as B_{lim}*exp(1.64*σ), where σ is assessment standard error in terminal year. Two options for σ were available: one from parametric bootstrap (σ =0.15) and standard value of 0.2. Due to uncertainty in catch composition of herring and sprat more precautionary value of 0.2 was used as σ. That produced B_{pa} of 570 kt (rounded from 574 kt) and $B_{trigger}$ was accepted at the same level.

Segmented regression was not considered a good basis for B_{lim} as its breakpoint (SSB of ca. 800 kt) generally separates data into two groups: one representing period before 1990 (above average and generally high cod biomass and thus high predation mortality of sprat) and the other from 1990 onwards (below average and mostly low cod biomass and low sprat predation mortality) (Figure 4.2). The recruitment for data from both periods show lack of dependence on cod SSB, and for period before 1990 recruitment shows even a declining trend with increasing sprat SSB. In such a case B_{lim} could be defined as B_{loss}, but as even at low sprat biomass (period before 1990) it does not show decline with biomass, B_{loss} from that period should be considered as B_{lim} (ca. 200 kt). For precautionary reasons such a low B_{lim} was not accepted.

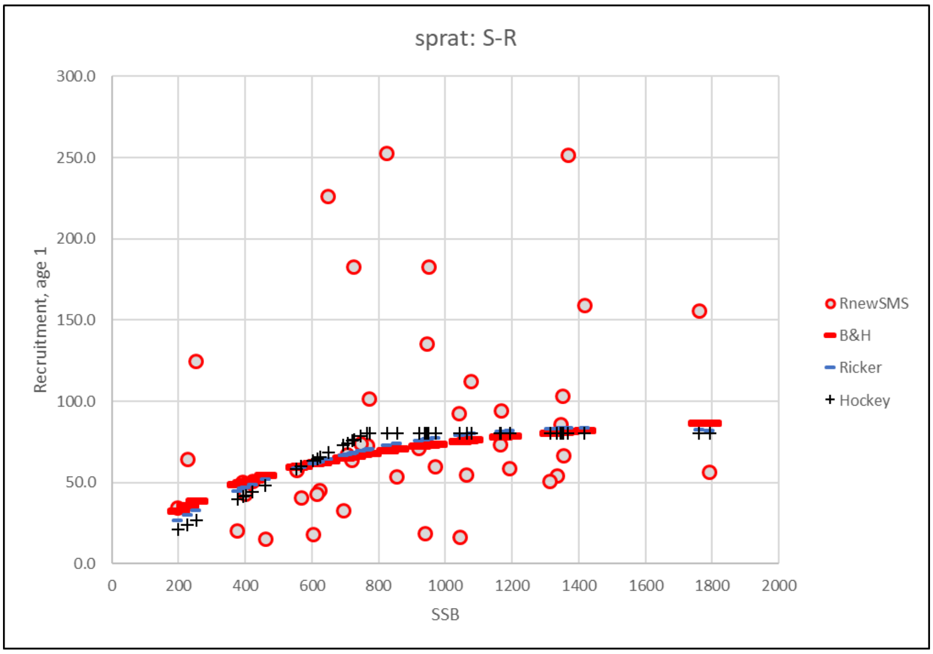

Figure 4.1. Stock–recruitment relationship for Baltic sprat; observations and fits of three models (Beverton and Holt, Ricker, and segmented regression).

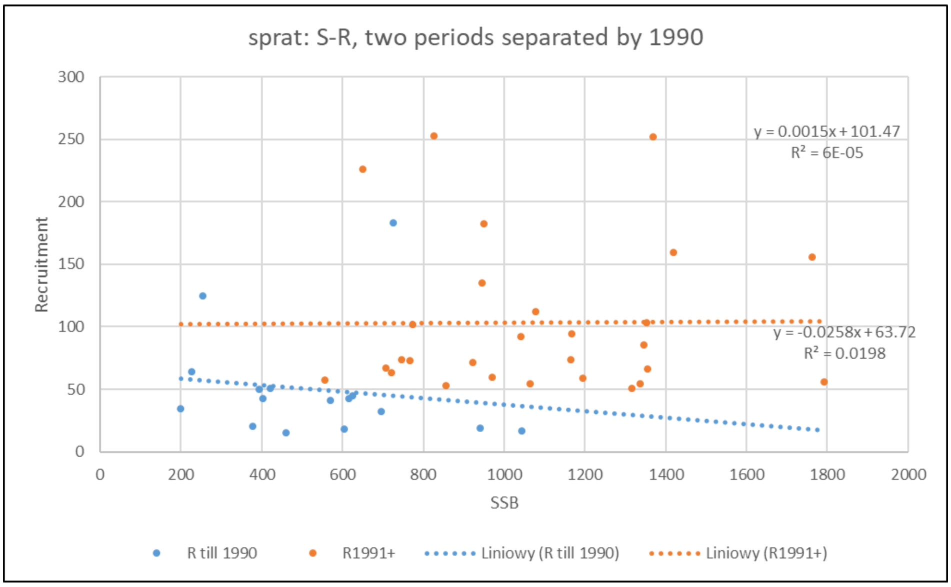

Figure 4.2. Stock–recruitment relationship for Baltic sprat separated into two periods: data before 1990 and data from 1990 onwards.

4.2 F_{MSY}, ranges and F_{pa} reference points

4.2.1 Choice of S–R relationship for the F_{MSY} simulations

As a first step, an analysis was made to examine which of the three, or combination of, S–R relationship best explained the relationship between SSB and recruitment of sprat. The analyses revealed that while the Beverton and Holt function explained 72% of the pattern, the segmented regression and the Ricker function explained only 23 and 4% respectively (Figure 4.3). Due to the low weight of the segmented regression and Ricker function these were not included further in the estimation of F_{MSY}.

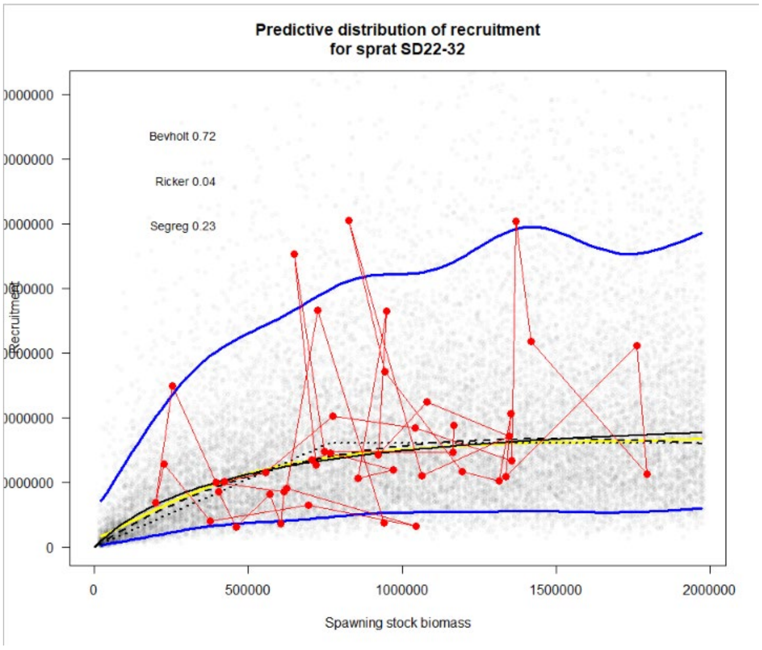

Figure 4.3. Explored S–R specifications. The numbers represent the weights of the different functions in explaining the S–R pattern of sprat.

4.2.2 F_{MSY}, ranges and F_{pa} reference points

For the F_{MSY} simulations the following year-ranges were used for biological parameters (weights, natural mortality) and fishing pattern:

Biological parameters: 2016–2018 (the last three years)

Fishing pattern: 2009–2018 (the last ten years)

B_{lim} was set to 410 000 t.

B_{pa} was set to 1.4* B_{lim} = 570 000 t and $B_{trigger}$ = B_{pa}.

As described above, the F_{MSY} simulations were run using the Beverton and Holt S–R function. When allowing the program to use the full range, and combinations of, bootstrap simulated a and b parameters in the Beverton and Holt function, the results presented unrealistically high catches at low fishing mortalities (Figure 4.4). The resulting F_{MSY} values were therefore not considered reliable. The extreme values of the parameters were thus removed (Figure 4.5) and the simulations run with this trimmed set of a and b parameters. A F_{MSY} simulation using the trimmed set of a and b parameters for the Beverton and Holt function resulted in a F_{MSY} of 0.31, with a range of 0.22–0.41 (Table 4.1a and b; Figure 4.6a and b). Note that as the $F_{MSYupper} > F_{0.05}$ it is capped to $F_{0.05}$ = 0.41.

If F_{MSY} has been restricted by F_{p05}, then $F_{MSYupper} = F_{MSY}$.

F_{lim} was estimated using the trimmed set of a and b parameters for the Beverton and Holt function as used in the F_{MSY} simulations. This resulted in a F_{lim} of 0.63, which corresponds to F_{pa} of 0.45 (F_{lim}*(exp(-1.645*0.2)). These values were accepted as final estimates of F_{lim} and F_{pa}.

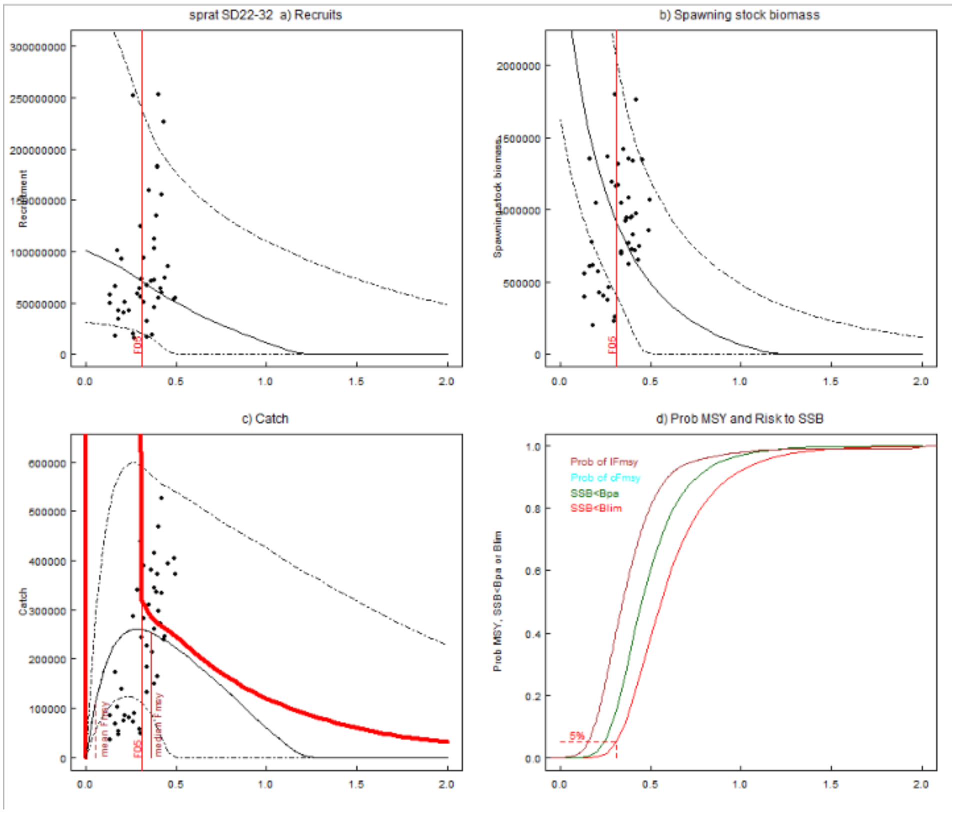

Figure 4.4. Results of the F_{lim} simulation using the full range of bootstrap simulated a and b parameters in the Beverton and Holt function.

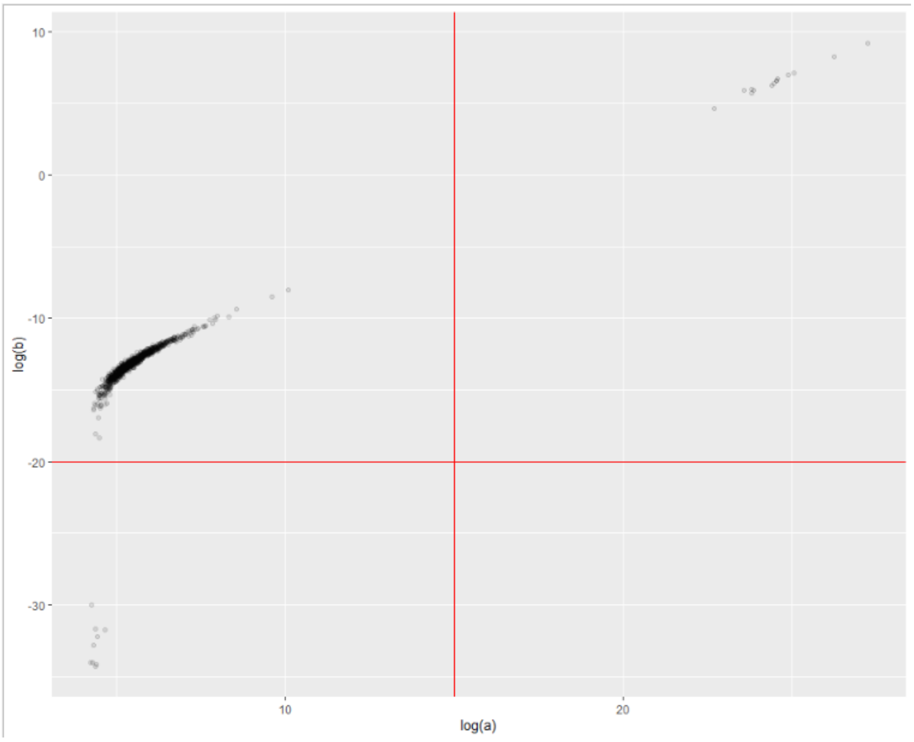

Figure 4.5. All bootstrap simulated a and b parameters in the Beverton and Holt function. The parameters within the left upper square are kept in the analyses while the other values are trimmed off.

Table 4.1a. Results of F_{MSY} simulations without advice rule.

FmsyMedianC	0.3115578
FmsylowerMedianC	0.2211055
FmsyupperMedianC	0.4321608
FmsyMedianL	0.3115578
FmsylowerMedianL	0.2211055
FmsyupperMedianL	0.4321608
F5percRiskBlim	0.3318311
Btrigger	0.0000000

Table 4.1b. Results of Fmsy simulations with advice rule.

FmsyMedianC	0.3115578
FmsylowerMedianC	0.2311558
FmsyupperMedianC	0.5226131
FmsyMedianL	0.3115578
FmsylowerMedianL	0.2311558
FmsyupperMedianL	0.5226131
F5percRiskBlim	0.4062513
Btrigger	574000.0000000

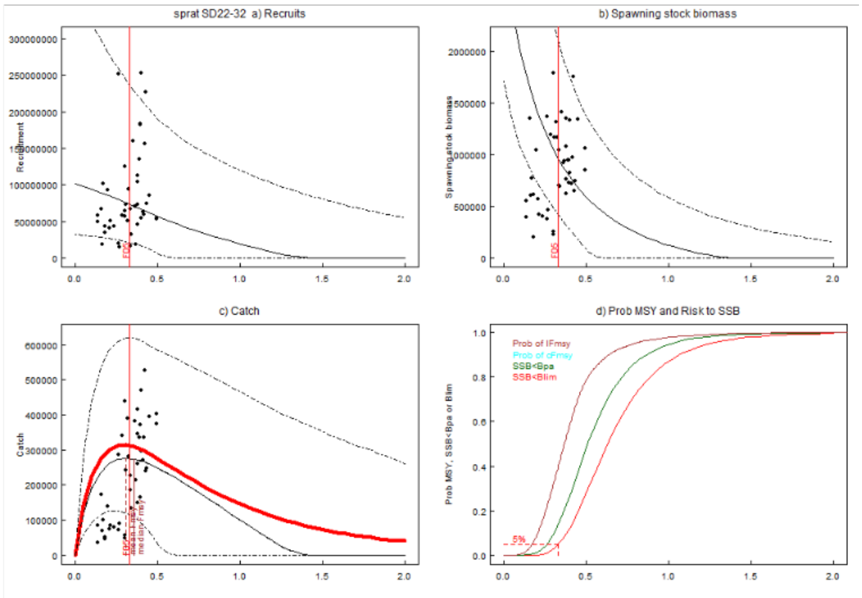

Figure 4.6a. F_{MSY} simulation without the advice rule.

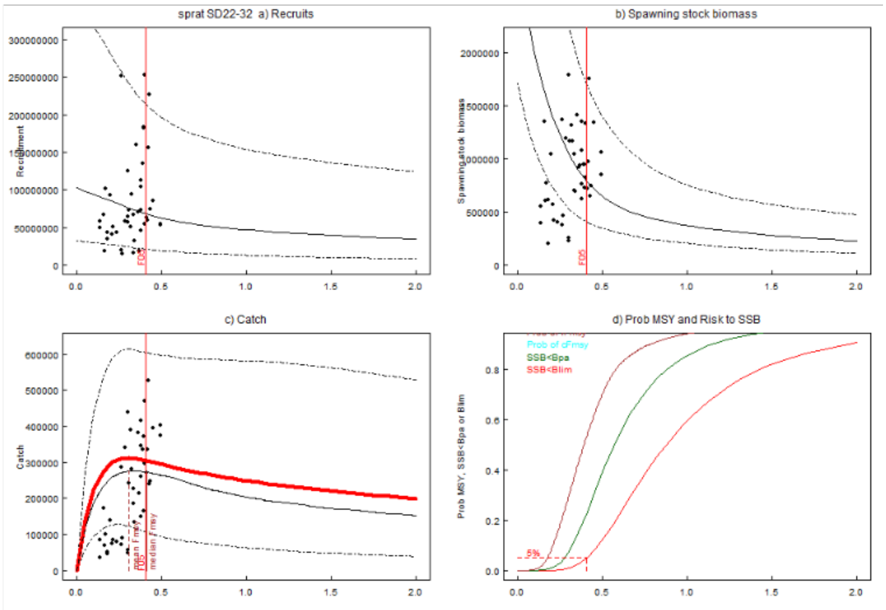

Figure 4.6b. F_{MSY} simulation with the advice rule.

Summary-New reference points

Reference Point	Value	Rationale
B_{lim}	410 000 t	The average SSB producing 50% of maximal recruitment from the Beverton and Holt S–R function (470 000 t) and from the Ricker S–R function (345 000t).
B_{pa}	570 000 t	1.4* B_{lim}
MSY $B_{trigger}$	570 000t	B_{pa}
F_{MSY}	0.31	Estimated by EqSim
$F_{MSYUpper}$	0.41	$F_{p0.5}$
$F_{MSYLower}$	0.22	Estimated by EqSim as the F at 95% of the landings of F_{MSY}
F_{lim}	0.63	Estimated by EqSim as the F with 50% probability of SSB being less than B_{lim}
F_{pa}	0.45	F_{lim} *(exp(-1.645*0.2))

Previous reference points – for comparison

Reference Point	Value	Rationale
B_{lim}	410 000 t	SSB producing 50% of maximal recruitment from the Beverton and Holt S–R function.
B_{pa}	570 000 t	1.4* B_{lim}
MSY $B_{trigger}$	570 000 t	B_{pa}
F_{MSY}	0.26	Annex I columns A and B in EU (2016)
$F_{MSYUpper}$	0.27	Consistent with ranges provided by ICES (2015), resulting in no more than 5% reduction of long-term yield compared with MSY (Annex I column B in EU (2016)
$F_{MSYLower}$	0.19	Consistent with ranges provided by ICES (2015), resulting in no more than 5% reduction of long-term yield compared with MSY (Annex I column A in EU (2016)
F_{lim}	0.39	Consistent with B_{lim}, ICES (2013)
F_{pa}	0.32	Consistent with B_{pa}, ICES (2013)

From these tables it is seen that the biomass reference points were unchanged, while all fishing mortality reference points increased from the values previously used.

5 Herring reference points

5.1 B_{lim} and B_{pa}

Analyses were conducted to search for an appropriate breakpoint in S–R relationship (shown in Figure 5.1). This was done by fitting segmented regressions at different SSB levels at intervals of 50 kt, to look at which level of SSB the lowest AIC of the fit would be obtained. The results indicated no clear breakpoint in these data (Figure 5.2). Lowest AIC values were obtained for breakpoints around 1 350 000 t, which is entirely driven by the high historical SSB values, and not considered appropriate for B_{lim}.

This suggest a Type 5 S–R relationship showing no clear evidence for impaired recruitment, in which case B_{lim}= B_{loss} is appropriate.

Type 1 could also be considered, corresponding to stocks with occasional large year-classes (as has been the case for herring in later decades). In this case B_{lim} would correspond to the lowest SSB where large recruitment is observed. The lowest SSB that has given rise to above average recruitment is the value from 2002. The SSB in 2002 happens to correspond also to B_{loss}, thus supporting B_{lim} =SSB 2002= 328 962 t. This corresponds to B_{pa} at 460 547 t (1.4* B_{lim}). These values were rounded to B_{lim}= 330 000 t and B_{pa}= 460 000 t.

$B_{trigger}$ is set equal to B_{pa}.

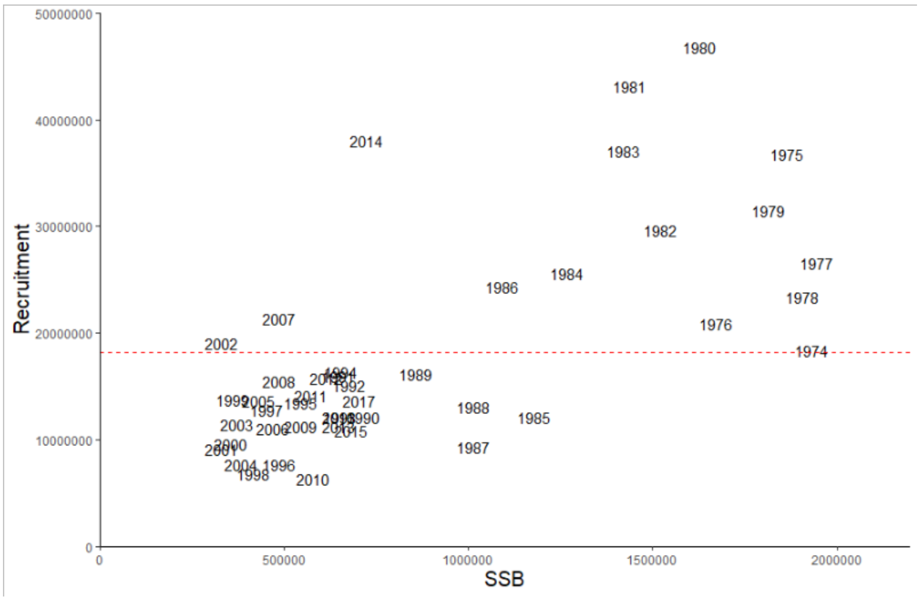

Figure 5.1. S–R relationship from the revised assessment. Red horizontal line shows average R in the time-series.

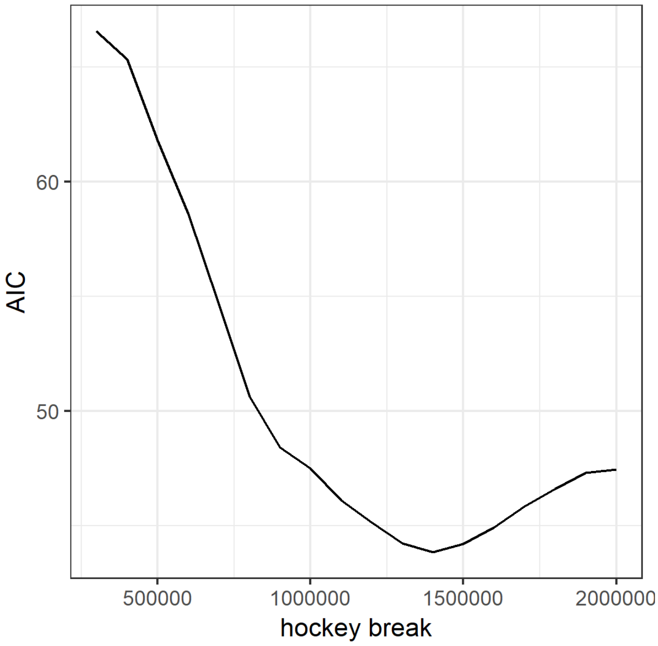

Figure 5.2. AIC results from fitting of segmented regression at different SSB levels.

5.2 F_{MSY}, ranges and F_{pa} reference points

5.2.1 Sensitivity of F_{MSY} to S–R specification

F_{MSY} was explored both with full and truncated (excl. years before 1990) time-series, although truncating the time-series is not well justified with present biological knowledge.

In all of the runs the following year ranges were used for biological parameters (weights, natural mortality) and fishing pattern:

Biological parameters: 2016–2018

Fishing pattern: 2008–2018

Sensitivity of F_{MSY} to different S–R specifications was tested, applying S–R specifications shown in Table 5.1.

The analyses revealed a high impact of S–R specification on estimated F_{MSY}, with the values varying from 0.10 to 0.35, between the scenarios. The purpose of this exercise is only to demonstrate that careful consideration should be given to, which S–R specification would be most appropriate to use, while all the explored scenarios are not considered equally appropriate.

Table 5.1. S–R scenarios explored.

Scenario ID	SR	F_{MSY}
1	TS from 1990, comb BH, Ricker, Segreg	0.16
2	TS from 1990, breakpoint at B_{lim}	0.35
3	TS from 1990, BH	0.16
4	TS from 1974, comb BH, Ricker, Segreg	0.11
5	TS from 1974, breakpoint at B_{lim}	0.31
6	TS from 1974, BH	0.10
7	TS from 1974, breakpoint at average SSB	0.21

5.2.2 Final F_{MSY} and ranges

The scenarios using truncated time-series are not considered appropriate due to lack of strong scientific basis for truncating the time-series.

In terms of S–R function, the explored scenarios showed that the lowest F_{MSY} values are derived from scenarios, where R is expected to continue to increase at higher stock sizes (i.e. scenarios where B–H relationship is dominant). Opposite, the highest F_{MSY} values are estimated when R is set to be entirely independent of SSB, e.g. set constant for all SSB values from above the lowest observed (which corresponds to B_{lim}).

None of these extreme assumptions can be considered well justified with the existing biological knowledge on the stock. The updated natural mortality estimates point in the direction of higher R at higher SSBs. Thus, it would be contradictory to this new information on M to apply constant low R at all levels of SSB from B_{loss} onwards. On the other hand, it is also highly unlikely that recruitment is an ever-increasing function of SSB (e.g. as B–H function indicates). Moreover, doubts have been raised whether the high recruitments seen in the past (1970s–1980s) are realistic at present ecological conditions in the Baltic Sea, which has however not been explicitly investigated for herring.

For these reasons, an intermediate option seems most appropriate. According to ICES Guidelines (p. 15), in cases where R seems to continue to increase with SSB (however is unrealistic to assume to continue), the change point of the segmented regression should be set at the average of all observed SSBs, corresponding to scenario 7 in Table 5.1, Figure 5.3. This option is considered a reasonable compromise between using the entire time-series of updated information, while not giving too high weight to the pre-1990s values.

This results in F_{MSY} at 0.21, with the ranges 0.15–0.26 (see Table 5.2). The results of Eqsim simulations are shown in Figure 5.4 and 5.5. $B_{trigger}$ has been set equal to B_{pa} in these simulations.

5.2.3 F_{lim} and F_{pa}

F_{lim} was estimated as F corresponding to 50% probability for SSB >B_{lim}. In these simulations estimating F_{lim}, breakpoint in S–R relationship was set to B_{lim}, according to ICES guidelines.

This resulted in F_{lim} estimate at 0.59, which corresponds to F_{pa} at 0.43 (F_{lim}*(exp(-1.645*0.2))

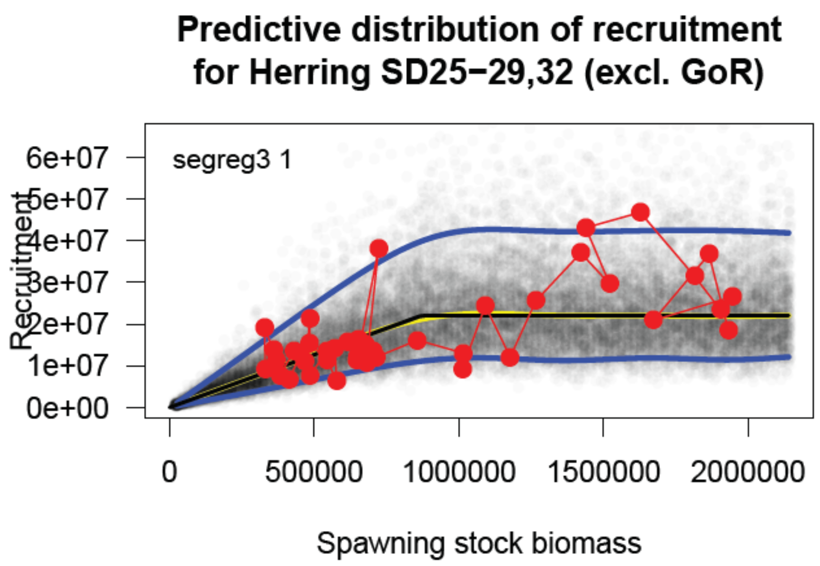

Figure 5.3. S–R relationship applied in F_{MSY} estimation.

Table 5.2a. Results of F_{MSY} simulations without advice rule.

```
FmsyMedianC      0.2110553
FmsylowerMedianC 0.1507538
FmsyupperMedianC 0.2613065
FmsyMedianL      0.2110553
FmsylowerMedianL 0.1507538
FmsyupperMedianL 0.2613065
F5percRiskBlim   0.2632287
Btrigger         0.0000000
```

Table 5.2b. Results of F_{MSY} simulations with advice rule.

```
FmsyMedianC      2.110553e-01
FmsylowerMedianC 1.507538e-01
FmsyupperMedianC 2.613065e-01
FmsyMedianL      2.110553e-01
FmsylowerMedianL 1.507538e-01
FmsyupperMedianL 2.613065e-01
F5percRiskBlim   3.244955e-01
Btrigger         4.571183e+05
```

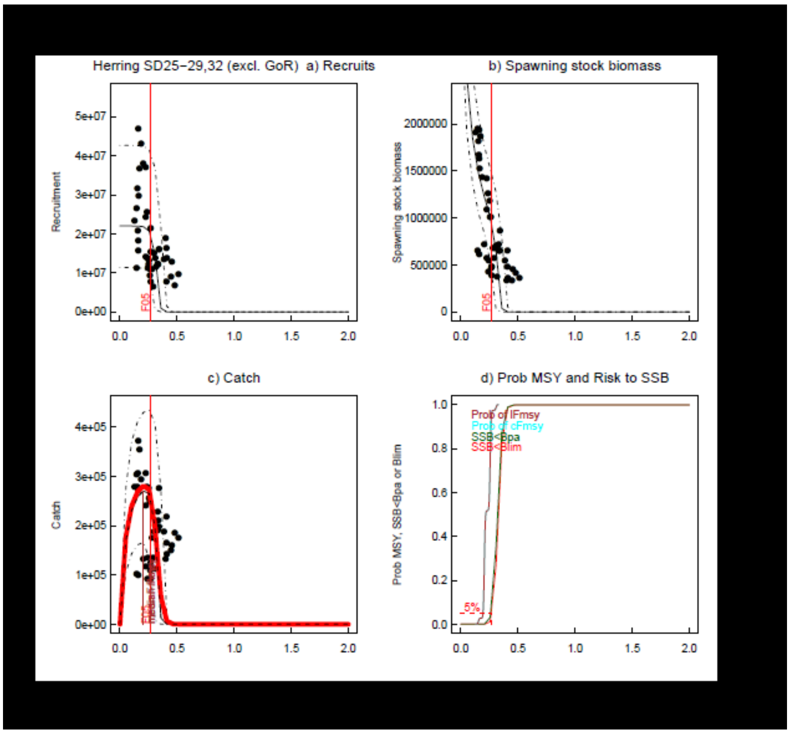

Figure 5.4. Results of initial F_{MSY} simulation (without Advice rule).

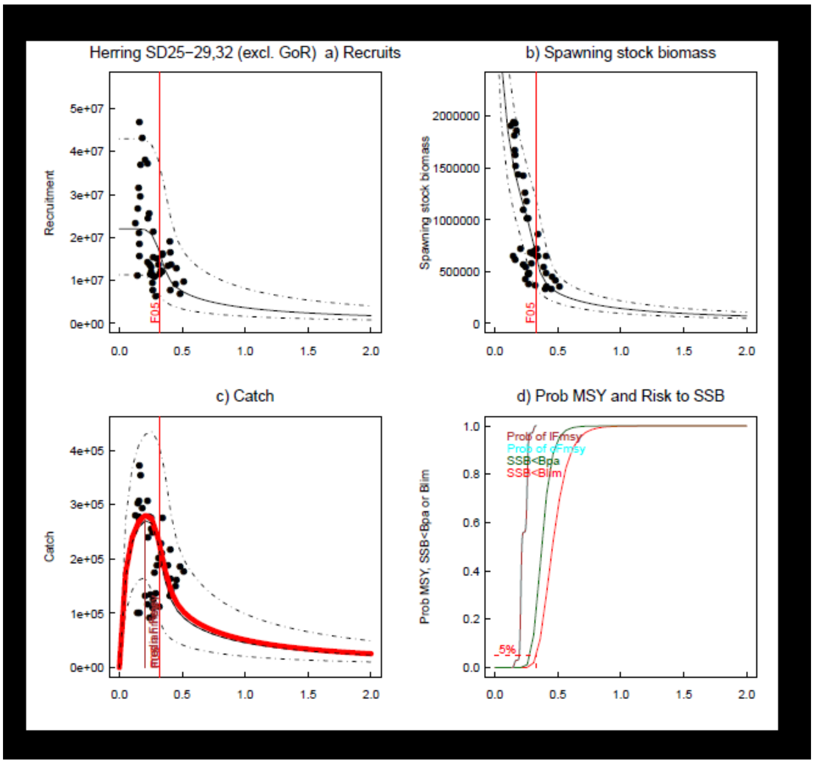

Figure 5.5. Results of F_{MSY} simulation with Advice rule.

Summary-new reference points

Reference Point	Value	Rationale
B_{lim}	330 000 t	The lowest SSB that has given rise to above average recruitment, i.e. year 2002. (The SSB in 2002 also happens to correspond also to B_{loss})
B_{pa}	460 000 t	1.4* B_{lim}
MSY $B_{trigger}$	460 000 t	B_{pa}
F_{MSY}	0.21	Estimated by EqSim
$F_{MSYUpper}$	0.26	Estimated by EqSim as the upper value of F at 95% of the landings of F_{MSY}
$F_{MSYLower}$	0.15	Estimated by EqSim as the lower value of F at 95% of the landings of F_{MSY}
F_{lim}	0.59	Estimated by EqSim as the F with 50% probability of SSB being less than B_{lim}
F_{pa}	0.43	F_{lim} *(exp(-1.645*0.2))

Previous reference points-for comparison

Reference Point	Value	Rationale
B_{lim}	430 000t	B_{loss}
B_{pa}	600 000 t	1.4* B_{lim}
MSY $B_{trigger}$	600 000 t	B_{pa}
F_{MSY}	0.22	Annex I columns A and B in EU (2016)
$F_{MSYUpper}$	0.28	Consistent with ranges provided by ICES (2015), resulting in no more than 5% reduction of long-term yield compared with MSY (Annex I column B in EU (2016)
$F_{MSYLower}$	0.16	Consistent with ranges provided by ICES (2015), resulting in no more than 5% reduction of long-term yield compared with MSY (Annex I column A in EU (2016)
F_{lim}	0.52	Consistent with B_{lim}, ICES (2013)
F_{pa}	0.41	Consistent with B_{pa}, ICES (2013)

From these tables it is seen that the biomass reference points were lowered by about 25%. F_{MSY} and the corresponding range were practically unchanged, while F_{lim} and F_{pa} increased slightly.

6 Comments from reviewers

6.1 Comments from Simon Fischer

The Inter-Benchmark Process (IBP) on BAltic Sprat (*Sprattus sprattus*) and Herring (*Clupea harengus*) (IBPBASH) was held in March 2020 remotely via WebEx meetings (6, 16 and 24 March) to include new estimates of natural mortality (M) from the WGSAM 2019 Baltic Sea SMS keyrun (ICES, 2019b) for the following two stocks:

- Sprat (*Sprattus sprattus*) in subdivisions 22–32 (Baltic Sea);
- Herring (*Clupea harengus*) in subdivisions 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea).

The following issues were addressed:

Natural mortality

Both stocks use the XSA model. Natural mortality was used from an SMS model run of WGSAM (ICES, 2012) which provided values up to 2011. In subsequent years, due to a lack of new SMS keyruns, natural mortality was estimated based on a linear regression of the average M (M1+M2) against the Baltic Sea cod SSB (the main predator). In the absence of an agreed ICES stock assessment SSB estimate for cod, cod SSB was approximated with an SSB survey index. In 2019, WGSAM produced a new SMS keyrun with M estimates for sprat and herring (ICES, 2019b). Historical M estimates were largely consistent with previous estimates and IBPBASH decided to use unsmoothed M values from WGSAM for the entire period.

For future years, until a new SMS keyrun will be available, it was decided to retain the previous approach and approximate M with a linear regression against cod SSB, which is currently available from an ICES stock assessment and the regression appears reasonable.

Stock assessments

Stock assessments were conducted with data and configurations identical to last year's working group (ICES, 2019a), apart from the new M values. Stock assessment results and trends were largely consistent with previous estimates. For herring, general SSB trend over the historical period remained unchanged, however, SSB estimates deviate increasingly since the early 1980s compared to WGBFAS 2019, with the 2018 SSB being around 20% smaller. For sprat, the new M values caused only negligible changes in stock assessment results. Model diagnostics for both stocks were similar to previous runs and model diagnostics, residual patterns and retrospective analyses were reasonable and did not indicate major issues.

Reference points

New reference points were calculated for both stocks based on the new stock assessment results and following extensive trials of different parameterisations and discussions. The main issue appeared to be that none of the stocks showed a clear stock–recruitment relationship on which calculations can be reliably based.

For herring, all reference points were calculated following ICES guidelines (ICES, 2017) and the analyses used the entire time period available from the stock assessment. B_{lim} was set to lowest observed biomass B_{loss}, B_{pa} set to $1.4B_{lim}$ and MSY $B_{trigger}$ to B_{pa}. F_{MSY} and ranges were estimated with EqSim (https://github.com/ices-tools-prod/msy) and following the guidelines for stocks where recruitment appears to increase continuously, a hockey-stick stock–recruitment model with the breakpoint manually set to the average SSB was used. F_{lim} was then derived by simulating a hockey-stick model with breakpoint at B_{lim} and determining the F which gave a 50% probability that SSB is above B_{lim}; F_{pa} was approximated as $F_{lim}/1.4$.

For sprat, the calculation of B_{lim} deviated from the ICES guidelines and B_{lim} was selected as the average of the SSB that produces 50% of the maximum recruitment from fitted stock–recruitment models (Beverton–Holt and Ricker), which is the same approach used when the reference points were estimated last (ICES, 2013). B_{pa} was then calculated as $B_{pa}=1.4B_{lim}$ and MSY $B_{trigger}$ set to B_{pa}. F_{MSY} and ranges were estimated with EqSim using a Beverton–Holt stock–recruitment model. During the bootstrapping process in EqSim, the Beverton–Holt model parameters α and β had to be manually constrained in order to avoid large catches at low fishing mortality. $F_{msyUpper}$ was limited to $F_{p0.5}$. Two values for F_{lim} were proposed; the first following the ICES guidelines and using B_{lim} as breakpoint for the forced hockey-stick stock–recruitment model and an alternative where the Beverton–Holt model from the estimation of F_{MSY} was used; F_{pa} was then calculated as $F_{lim}/1.4$.

The EqSim software conducts a stochastic simulation. For IBPBash, random number seeds were not set or recorded prior to running the simulation and therefore exact reproducibility cannot be guaranteed.

Conclusion

The reviewers were requested to submit their comments prior to the compilation of the report and therefore are rather a review of the working documents, presentations and discussion during the inter-benchmark process meetings and do not necessarily correspond to the final report.

The natural mortality values from the recent SMS keyrun (WGSAM; ICES, 2019a) and their implementation into the stock assessments of Baltic sprat and herring as parameterized during this benchmark can be considered appropriate and provide a novel source of data not previously available (ToR a).

ToR b requested updates of the stock annex. Updated stock annexes were not provided prior to the date by which the review was requested to be submitted and could therefore not be reviewed.

Finally, the procedure for estimation of updated reference points (ToR c) largely followed ICES guidelines or previous approaches for these stocks and appears reasonable.

Further recommendations

Only for one of the two benchmarked stocks (herring), the stock assessment input and configuration files were made available on the ICES SharePoint before the first WebEx meeting which allowed participants and reviewers to reproduce the stock assessment with the old and new natural mortality values and verify the analysis. For future workshops, it is recommended to make all files and configurations which are required to run the stock assessment available for all stocks; the ICES Transparent Assessment Framework (TAF) would be obvious choice. Furthermore, providing input data and scripts used for the calculations of the reference points, including random number seeds, would be useful and ensure reproducibility.

This inter-benchmark dealt with the inclusion of novel natural mortality values into the existing stock assessment configuration. Future benchmarks could review the appropriateness of data sources, model configurations and might also consider alternative assessment models (SAM is already considered regularly during WGBFAS).

6.2 Comments from Marc Taylor

In response to new estimates of natural mortality (M) from the WGSAM 2019 Baltic Sea SMS keyrun, the Inter-Benchmark Process (IBP) was opened for the following stocks: Sprat (*Sprattus sprattus*) in subdivisions 22–32 (Baltic Sea) and Herring (*Clupea harengus*) in subdivisions 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea). The ToRs of the IBP were to evaluate the appropriateness of the updated M values, update the stock annexes and re-examine and reference points.

In both stocks, the updated M values were deemed acceptable and would be used in future assessments. Other settings used in the XSA assessment remain unchanged.

For both stocks, variations in predation mortality, M2, are assumed to be mainly dependent on cod predation, which is treated as an "other predator" in the SMS model. Due to age-reading problems for cod in the eastern Baltic, ICES now applies and age–length based analytical assessment with the Stock Synthesis model (SS3). Due to the intermittent nature of SMS keyruns, intermediate assessment years have estimated M for herring and sprat assessments via linear regressions between cod SSB indices and mean M (response variable, across ages). These regressions have been updated and their continued use was deemed acceptable. An alternate relationship between average M and biomass of cod >35 cm was tested, but SSB was shown to be a better predictor covariate.

Despite similarities in historical M trends, magnitudes changed to differing degrees by age classes which resulted in changes to various assessment outputs (e.g. F, SSB, recruitment), and reference points were thus updated. A majority of the discussion focussed on the procedures to follow for these updates, which are elaborated below.

Sprat

Although exhibiting a similar temporal trend, the youngest age groups (ages 0–1) showed large increases in M. Catchability q estimates remained similar. Both Beverton–Holt and Ricker stock–recruitment relationships (SRRs) were found to fit the data better than segmented regression, and thus were used in defining B_{lim} (biomass leading to 50% of maximum observed, and average value of the two SRRs). The resulting B_{lim} was also deemed to be more conservative than that resulting from the segmented regression. B_{pa} was left to a more conservative, standard estimate of B_{pa}=1.4*B_{lim}, which seems appropriate given the assumed data quality.

The EqSim SRR permutation routine, used in the establishment of other reference points, also identified the typically-used segmented regression SRR as a poor model. The Beverton–Holt model most often best explained the permutated SRR pattern, and was thus chosen for further reference point definitions. Following updates in biological parameters and fishing patterns, the EqSim routine estimated equilibrium yields under variable F. Biological parameter resampling was done over a more recent time span of more current conditions (three years), while exploitation patterns were resampled from a longer ten-year period. The use of alternate settings for the SRR were well-founded and provided more conservative and precautionary reference points. Remaining reference point definitions followed the ICES guidelines.

Herring

Largest differences in M were observed for the beginning of the time-series (i.e. higher). The age 1 class showed higher overall estimates while ages 2 and above showed lower estimates. The unsmoothed time-series from SMS were used, as with sprat.

The choice of "Stock type" for the reference point guidelines was complicated by competing evidence for more than one type, as well as whether a possible regime shift should don-weight the importance of historical values. Evidence was given for both Type 5 SRR relationship (showing no clear evidence for impaired recruitment) as well as Type 1 (Spasmodic stocks, stocks with occasional large year classes). Both indicate that B_{lim} = B_{loss}. As with sprat, B_{pa} = 1.4*B_{lim} was assumed.

The EqSim SRR permutation routine tried several SRR configurations; including segmented regression, Beverton–Holt, and combined (Beverton–Holt, Ricker, segmented regression). Similar to sprat, temporal ranges used for resampling of biological parameters and fishing pattern were 3- and 10-years, respectively. Large differences in resulting reference points depended primarily on the assumptions of the SRR, although shortening the time extent was also tested, with similar trends in affect to F_{MSY}. Specifically, a segmented regression SRR that used the previously defined B_{lim}=B_{loss} resulted in a much higher F_{MSY} than that produced with a combined or Beverton–Holt SRR. Finally, a B_{lim} equalling the average of historical SSBs was justified according to the guidelines in cases where recruitment is observed to increase with SSB but without evidence to suggest this will not continue. This final alteration to B_{lim} resulted in an intermediate F_{MSY} = 0.21 when combined with a segmented regression.

Conclusions

The resulting IBP was conducted largely according to ICES guidelines. Deviations were well-justified and openly-discussed among all participants and reviewers. TORs a & c were evaluated by this reviewer and found to have been adequately addressed and reported. As mentioned by Simon Fisher, TOR b (re: updating stock annexes) was not available to the reviewers at the time of review.

Further suggestions

Both stocks show poor relationship between SSB and recruitment, and lack a clearly defined SSB leading to impaired recruitment (i.e. B_{lim}). Some discussion centred on whether there has been a regime shift in recruitment success (i.e. for herring), which would justify truncating the time-series in the fitting of an SRR. The groups' preliminary conclusion was that the evidence is too weak to justify this currently, although some of the decisions made in the parameterisation of EqSim still hinted at this assumption; e.g. the more or less linear relationship between recruitment and SSB was assumed unlikely to be maintained in the future. A more in depth analysis of the evidence for a regime shift is warranted in the future benchmark (e.g. change-point analysis of recruitment success).

Evidence for significant linear relationships between cod SSB and predation mortality was clearly shown and justified the use of the relationship's predicted M for assessed years occurring before and updated SMS keyrun becomes available. Further exploration into the underlying model assumptions (i.e. Holling's functional response curve type) may make for a more direct approach in deriving M values. It should be investigated if such an output from SMS is available, which could be used in future predictions of M.

Both reviewers find the quality of the assessment appropriate to be used as basis for advice.

7 References

EU. 2016. Regulation (EU) 2016/1139 of the European Parliament and of the Council of 6 July 2016 establishing a multiannual plan for the stocks of cod, herring and sprat in the Baltic Sea and the fisheries exploiting those stocks, amending Council Regulation (EC) No. 2187/2005 and repealing Council Regulation (EC) No. 1098/2007. Official Journal of the European Union, L 191. 15 pp. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R1139&rid=1.

ICES. 2012. Report of the Working Group on Multispecies Assessment Methods (WGSAM), 22–26 October 2012, Venice, Italy. ICES CM 2012/SSGSUE:10. 145 pp.

ICES. 2013. Report of the Benchmark Workshop on Baltic Multispecies Assessments (WKBALT), 4–8 February 2013, Copenhagen, Denmark. ICES CM 2013/ACOM:43. 399 pp.

ICES. 2015. EU request to ICES to provide F_{MSY} ranges for selected North Sea and Baltic Sea stocks. *In* Report of the ICES Advisory Committee, 2015. ICES Advice 2015, Book 6, Section 6.2.3.1. 11 pp.

http://www.ices.dk/sites/pub/Publication%20Reports/Advice/2015/Special_Requests/EU_FMSY_ranges_for_selected_NS_and_BS_stocks.pdf.

ICES. 2017. ICES fisheries management reference points for category 1 and 2 stocks. In ICES Advice Technical Guidelines. ICES Advice 2017, pp. 1–19.

ICES. 2019a. Baltic Fisheries Assessment Working Group (WGBFAS). ICES Scientific Reports. 1:20. 653 pp.

ICES. 2019b. Working Group on Multispecies Assessment Methods (WGSAM). ICES Scientific Reports. 1:91. 320 pp.

Myers, R.A., Rosenberg, A.A., Mace, P.M., Barrowman, N., Restrepo, V.R. 1994. In search for threshold of recruitment overfishing. ICES J. mar. Sci., 51:191–205.

Annex 1: List of participants

NAME	ADDRESS	COUNTRY	E-MAIL
Mikaela Bergenius	Department of Aquatic Resource Swedish University of Agricultural Sciences	Sweden	Mikaela.bergenius@slu.se
Bjarte Bogstad Chair	Institute of Marine Research	Norway	Bjarte.bogstad@hi.no
Margit Eero	DTU-Aqua	Denmark	mee@aqua.dtu.dk
Ruth Fernandez	ICES	Denmark	Ruth.fernandez@ices.dk
Simon Fischer Reviewer	Lowestoft Laboratory Cefas	UK	simon.fischer@cefas.co.uk
Julita Gutkowska	National Marine Fisheries Research Institute	Poland	jgutkowska@mir.gdynia.pl
Tomas Gröhsler	Thünen-Institute of Baltic Sea Fisheries	Germany	Tomas.groehsler@thuenen.de
Jan Horbowy	National Marine Fisheries Research Institute	Poland	horbowy@mir.gdynia.pl
Olavi Kaljuste	Institute of Coastal Research Swedish University of Agricultural Sciences	Sweden	Olavi.kaljuste@slu.se
Marc Taylor Reviewer	Thunen Institute	Germany	Marc.taylor@thuenen.de

Annex 2: Stock annexes

The table below provides an overview of the stock annexes updated at IBPBash. Stock Annexes for other stocks are available on the ICES website Library under the Publication Type "Stock Annexes". Use the search facility to find a particular Stock Annex, refining your search in the left-hand column to include the *year*, *ecoregion*, *species,* and *acronym* of the relevant ICES expert group.

Stock ID	Stock name	Last up-dated	Link
Her.27.25–2932_SA	Herring (*Clupea harengus*) in subdivisions 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea)	March 2020	Herring in SD 25–29 and 32
Spr.27.22–32_SA	Sprat (*Sprattus sprattus*) in Subdivisions 22–32 (Baltic Sea)	March 2020	Baltic Sea Sprat