# WORKING GROUP ON MULTISPECIES ASSESSMENT METHODS (WGSAM; outputs from 2022 meeting)

VOLUME 5 | ISSUE 12

ICES SCIENTIFIC REPORTS

RAPPORTS
SCIENTIFIQUES DU CIEM

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

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

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

ISSN number: 2618-1371

ICES Scientific Reports

Volume 5 | Issue 12

# WORKING GROUP ON MULTISPECIES ASSESSMENT METHODS (WGSAM; outputs from 2022 meeting)

## Recommended format for purpose of citation:

ICES. 2023. Working Group on Multispecies Assessment Methods (WGSAM; outputs from 2022 meeting).
ICES Scientific Reports. 5:12. 233 pp. https://doi.org/10.17895/ices.pub.22087292

## Editors

Valerio Bartolino • Michael Spence

## Authors

Alan Baudron • Alex Holdgate • Alfonso Perez • Andrea Belgrano • Andy Beet • Carolyn McKeon • Chris Legault • Christina Perez • Christopher Griffiths • Floor Soudijn • Francisco De Castro • Gavin Fay • Gustav Delius • Howard Townsend • Jacob Bentley • James Martindale • Jamie Lentin • Jamie Tam • Kiersten Curti • Liesa Celie • Maria Ching Villanueva • Marianne Robert • Marie Savina Rolland • Maxwell Grezlik • Michael Spence • Michael Thomson • Mikaela Potier • Miriam Puts • Morten Vinther • Robert Gamble • Robert Thorpe • Ron Klasky • Roweena Patel • Sarah Gaichas • Sean Lucey • Sigrid Lehuta • Thomas Del Santo O'Neill • Valerio Bartolino • Vanessa Trijoulet

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

# Contents

# i Executive summary

The Working Group on Multispecies Assessment Methods (WGSAM) aims to advance the operational use of knowledge on predator-prey interactions for advice on fisheries and ecosystem management.

This report presents an update of the multispecies SMS keyrun model for the Baltic Sea including its review by the working group, and the review of three modelling frameworks for the Georges Bank marine ecosystem. The Baltic Sea keyrun provides updated estimates of cod predation mortality for the Baltic Sea sprat and central Baltic herring stocks. The model integrates fishery and survey data on the two clupeids and makes extensive use of the cod stomach data (i.e. 64 000 stomachs are used as input to the model). Estimations of predation mortality are consistent with previous estimates and suitable for inclusion in the stock assessment of the two clupeid stocks. Predation remains low on all ages for both herring and sprat as a result of the low cod stock size. The large modelling framework for the Georges Bank comprises a multispecies production model, a multispecies length-based model and a mass-balanced ecosystem model. While none of the three models was at that stage to be evaluated as a keyrun, with further development they could complementarily and in conjunction contribute to the EBFM of the Georges Bank.

# ii Expert group information

| | |
|---|---|
| **Expert group name** | Working Group on Multispecies Assessment Methods (WGSAM) |
| **Expert group cycle** | Multiannual |
| **Year cycle started** | 2022 |
| **Reporting year in cycle** | 1/3 |
| **Chairs** | Valerio Bartolino, Sweden |
| | Michael Spence, UK |
| **Meeting venue and dates** | 10–14 October 2022, Woods Hole, USA (>40 participants) |

# 1 Baltic Sea SMS key-run review

## 1.1 SMS model

### 1.1.1 Is the model appropriate for the problem?

The SMS model will be used to provide natural mortality estimates by age and year as input to single species assessments of Eastern Baltic herring (Subdivisions 25–29 and 32 (excluding Gulf of Riga)) and Baltic sprat (Subdivisions 22–32). Natural mortality estimates are only used as input for the historic part and no forecast is needed. M estimates by age and quarter are a direct output of the model. However, an assumption is needed for residual mortalities M1 while the predation mortalities M2 are estimated (M = M1+M2). The model is able to provide estimates for the years 1974 to 2021.

The eastern Baltic cod occurs in the SD24–32 with largest abundances in SD25–29, and the western Baltic cod is found in the SD22–24. The distribution of the eastern Baltic stock overlaps entirely with the distribution of the central Baltic herring stock (SD25–29,32), while only to a large extent with the distribution of Baltic sprat which occurs throughout the SD22–32. The Baltic SMS key run includes the eastern Baltic cod but not the western Baltic stock, hence what referred to cod predation in the model is in fact limited to predation by the eastern Baltic stock.

Cod is the only predator in the model, feeding on herring sprat, and an ‘other food’ component. Cod is a constant predator in the SMS model due to age reading problems with the stock. For some years, only length-based input such as catch at length have been provided for the assessment of the cod stock. Such input cannot be handled in SMS. Therefore, numbers at age derived from the Eastern Baltic cod assessment are used as known input without error since the 2019 key run.

Overall, the model is appropriate to provide information on natural mortalities as input for the assessments of central Baltic herring and Baltic sprat. However, results depend to a larger extent on the input from the Eastern Baltic cod assessment. Since these are assumed to be known without error, SMS uncertainty estimates around M are likely underestimated and cannot be considered informative.

### 1.1.2 Is the scientific basis of the model sound?

The SMS model is an established and reviewed model that has previously been applied in the Baltic Sea to provide inputs for assessments of commercially exploited stocks (Baltic sprat and herring). It is also applied for the North Sea foodweb (ICES WGSAM 2014, 2017, 2020) to provide input natural mortalities for the assessments of many commercially exploited stocks in the North Sea (e.g., North Sea cod and herring). The Baltic SMS model has been reviewed by ICES in several occasions (ICES 2012a, 2012b, 2019).

### 1.1.3 Is the input data quality and parameterization sufficient for the problem?

Model setup

For herring and sprat, the data (1974–2021), in quarterly time steps, used as input to SMS consists of: catch at age, proportion landed, mean weight at age in the catch and the stock, proportion mature, proportion mortality before spawning, natural mortality at age, survey effort and survey catch at age (Lewy and Vinther, 2004). For cod, the data used as input to SMS consists of: cod numbers and mean weight per 2-cm length class from the ICES WGBFAS, Stock-Synthesis 3 (SS3) assessment from 1974–2021 and Latvian cod stomach data from 1974–2014. SMS is fitted to the Cod stomach data and the catch at age and survey indices for herring and sprat. SMS estimates values for F, SSB, recruitment and M2 (i.e., cod predation mortality) for herring and sprat. In total, 300+ parameter values are estimated by the model.

Data quality

SMS uses the same data as used for input to the single species assessments of herring and sprat in the Baltic Sea (catch at age, mean weights, proportion mature, survey indices). These data have been benchmarked and therefore no further review on these data has been carried out. However, while single species assessments start with age 1 as recruits, SMS starts with age 0. Given that predation mortalities are important for age 0, this is understandable. But e.g., the mean weight at age is highly uncertain for the 0 group. WGBFAS assumes that mean weight at age in the sea is the same as mean weight at age in the catch. In general, the assumption that mean fish weight in the catch equals the mean fish weight in the stock is a major assumption that can have an influence on model results. SMS uses smoothed quarterly values as input for the age 0 (Q3 & Q4) and age 1 (Q1 & Q2). Even though there is a large uncertainty around the input values for weight at age in the sea, especially for the youngest ages.

In addition to the data that is used for the single species assessments, SMS relies on cod stomach data. A new stomach data set for cod became available in recent years. Due to the better quality of the new stomach contents data, this data sets were applied as default in the key run of 2019 (ICES WGSAM, 2019). The stomachs have been collected mainly by the Latvian institute and provide detailed information based on individual stomachs. The dataset provides information for the period 1974–2014, stomach data from more recent years is not available. Overall, data from 64 000 stomachs were used as input for SMS. Stomach data are reported per individual cod with predator size classes in mm or cm. Although in general the eastern Baltic is well covered, sampling focuses around the Latvian coast and there are some gaps in space, especially in the 1990s. The main distribution areas for cod are the SD 25, 26 and 28 and stomachs from these areas are used for SMS both in this as in previous key runs. The three subdivisions have not simultaneously been sampled in all years.

The compilation procedure followed to parameterise stomach contents in the 2022 key-run was performed via the ‘FishStomach’ R package (https://github.com/MortenVinther/FishStomachs). This introduced some differences in terms of relative stomach contents (Annex 2, section 2.3.2) compared to the beta version of FishStomach used for the 2019 key-run, although these are minimal.

Consumption rates are calculated based on the cylinder gastric evacuation rate model from Andersen and Beyer (2005a, b). In order to consider recent changes in cod consumption rate, the relationship between average quarterly consumption rate and total length (a priori parametrized as $C=aL^b$ with C the average quarterly consumption rate and L total length) was estimated separately for three different periods (1974–1989, 1990–1999, 2000–2014).

Overall, the data quality is considered decent input to the model. That the model uses key input from the single species assessments can be seen as a strength because these data were already scrutinized during a full ICES benchmark process. It is a positive development that the stomach data compilation process has been documented in an R package, this should facilitate the reproducibility of the stomach data compilation process in the future, and avoid discrepancies in stomach data used in subsequent key runs.

<u>Assumptions and parameterization</u>

The parameterization of the diet selection sub-model is based on several assumptions. First, there is only one vulnerability parameter per interaction for the full model time-series. The assumption of constant vulnerability may be violated if e.g. the spatial predator-prey overlap changes. Given an assumed constant overlap, the implied Holling type II functional feeding response as used in SMS is well known to lead to instability when prey items become low in abundance and makes them vulnerable to extinction in the model. However, this is mainly an issue for forecasts when trying to make predictions outside the range of observations.

Another important assumption is a time constant biomass pool of "Other Food" that for the eastern Baltic cod includes a considerable proportion of benthic preys. If the availability of important Other Food prey items changes over time, this can lead to biased predictions of relative stomach contents and therefore predation mortalities. As shown in Figures 7 and 8 in the section 1.1.4 below, this could be an issue for this key-run. The current SMS key run is based on a log-normal size selection of the predators (see Annex 2).

Overall, the parameterization and assumptions are consistent with recent scientific knowledge. However, some of the assumptions regarding constant vulnerabilities and constant Other Food availability may be relaxed in future key-runs and further investigations are needed to fully utilize the extensive time-series of stomach data (also in relation to spatial patterns) to optimise the parameterisation of the diet selection sub-model.

## 1.1.4 Does model output compare well with observations?

The SMS key-run was able to estimate total catch over time sufficiently well for herring and sprat (Figure 1), and the main features of the time-series are well covered by the model. For herring, the catch-at-age residual plots by quarter do not show any discernible pattern (Figure 2). For sprat however, there are clear clusters of positive and negative residuals, except for the oldest age classes all show more positive residuals in the earlier part of the time-series, especially for Q3 (Figure 3). Catch-at-age residuals for sprat are also larger overall than for herring.

**Figure 1. Comparison between predicted and observed catch time-series.**

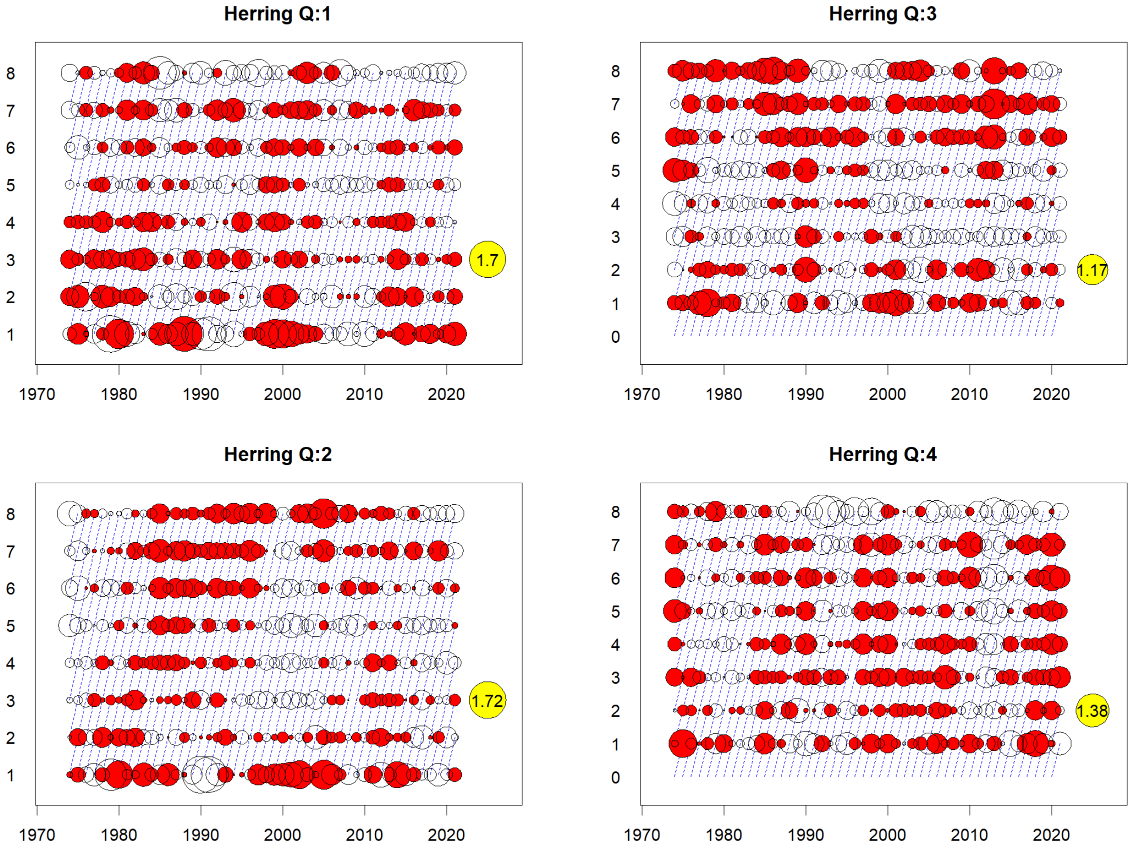


**Figure 2. Herring catch observations Residuals, log(Survey observed CPUE) - log(expected CPUE). Red is positive, White is negative. Q:9 means annual data. The yellow circle illustrates the size of the largest value as a reference.**

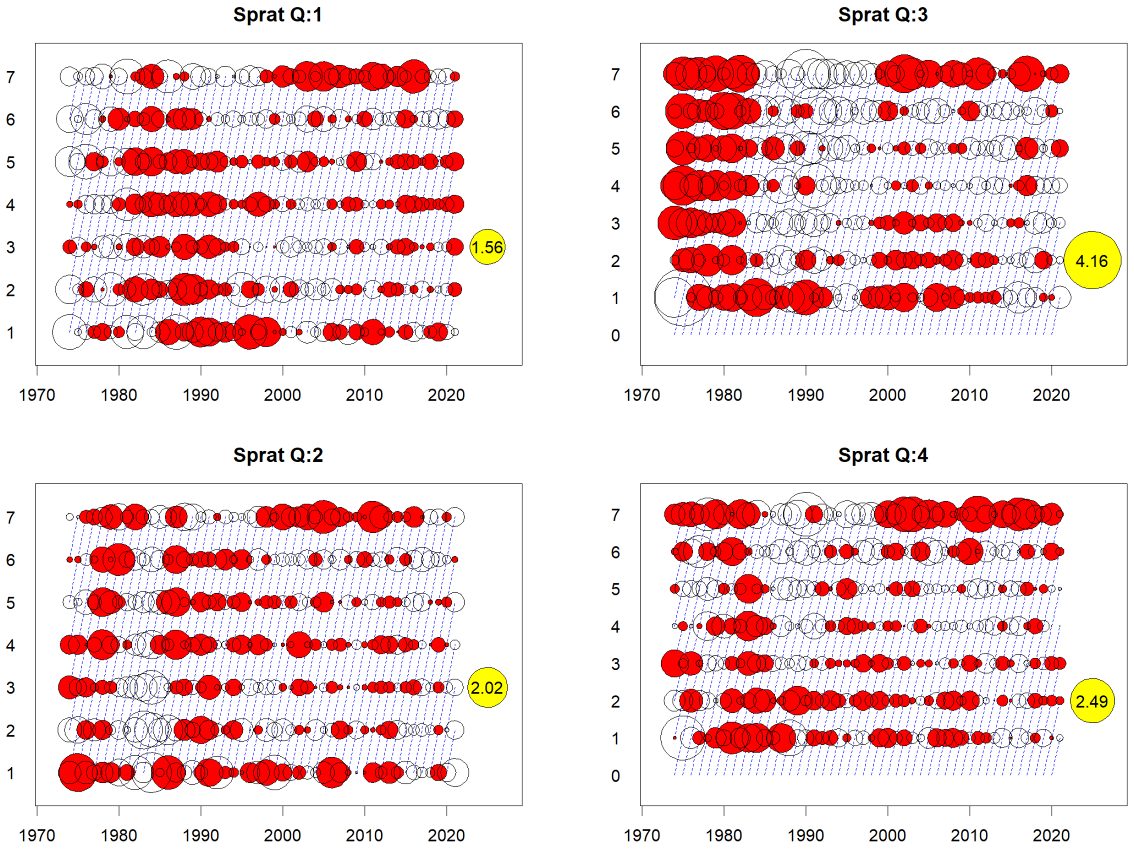


**Figure 3. Sprat catch observations Residuals, log(Survey observed CPUE) - log(expected CPUE). Red is positive, White is negative. Q:9 means annual data. The yellow circle illustrates the size of the largest value as a reference.**

The fit to survey data captures the main trends reasonably well for both herring (Figure 4) and sprat (Figure 5), with the exception of age 1 herring in the BIAS survey and age 1 sprat in the LAT RUS acoustic survey. There are clear "year effects" (consistent under- or overestimation in a given year), seen in the May acoustic surveys for both herring and sprat (and in the LAT RUS acoustic survey for sprat). However, both the BIAS survey for herring and the October acoustic survey for sprat do not show any discernible pattern (Figure 6). Overall, no sign of overfitting to a particular data source is apparent.

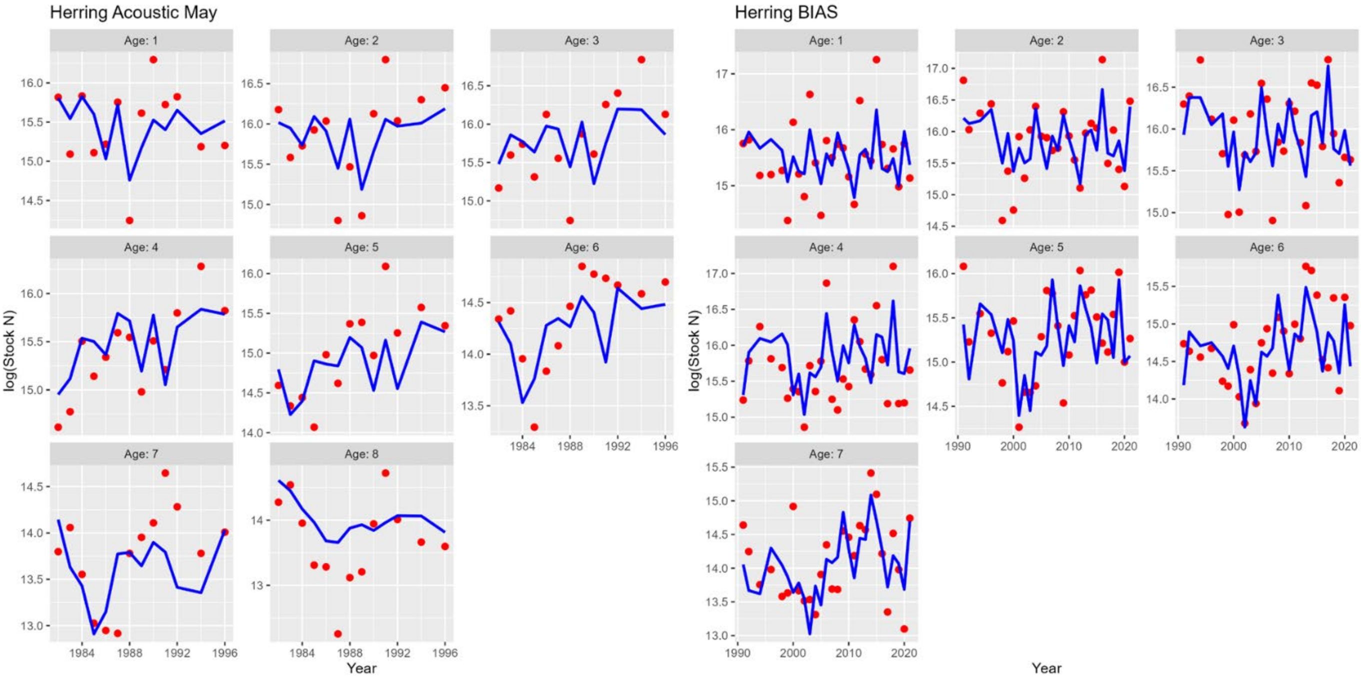


**Figure 4. Herring surveys observation residuals (log scale). The red dots are the observed values, the blue lines are the prediction.**

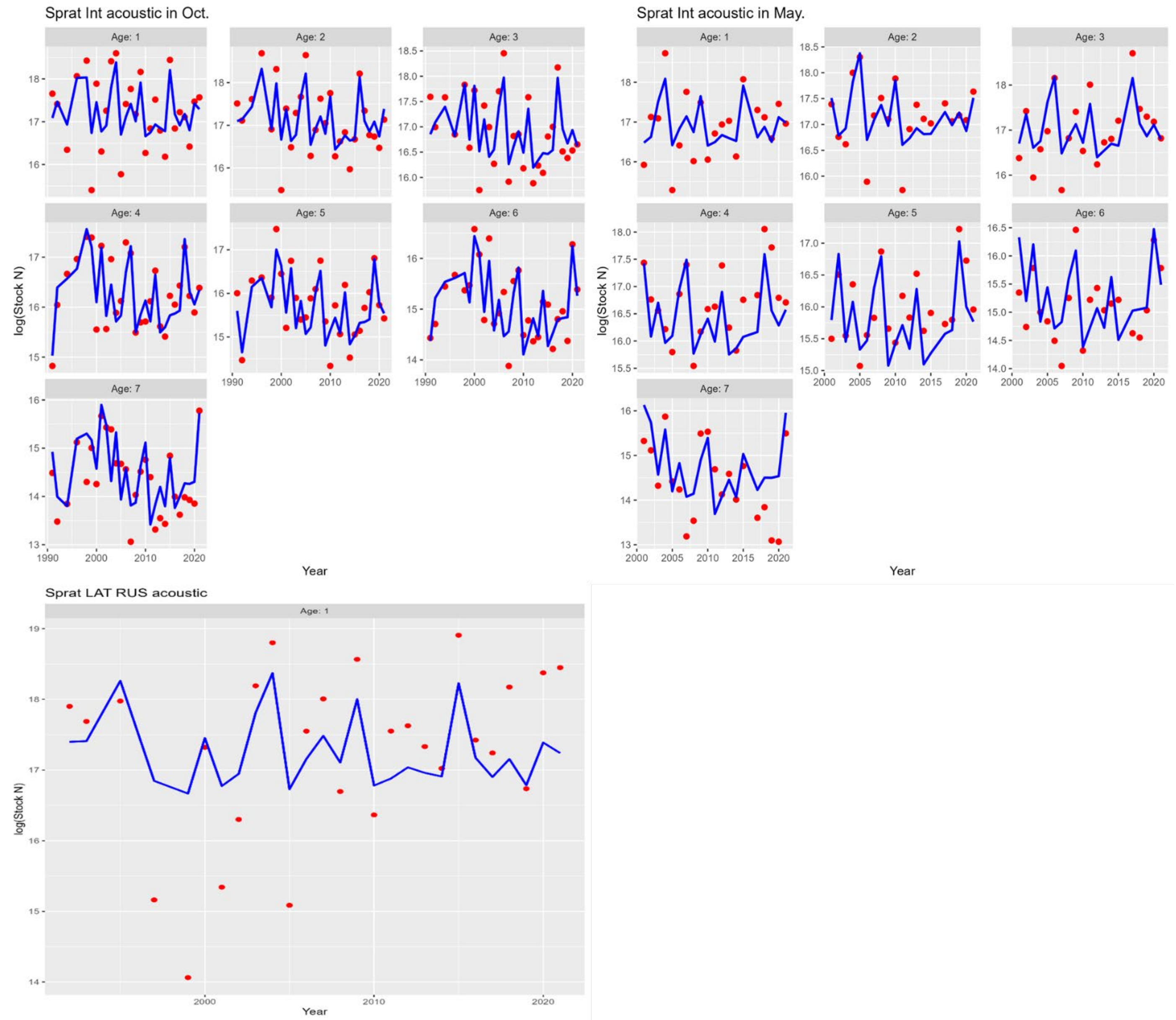


**Figure 5. Sprat surveys observation residuals (log scale). The red dots are the observed values, the blue lines are the prediction.**

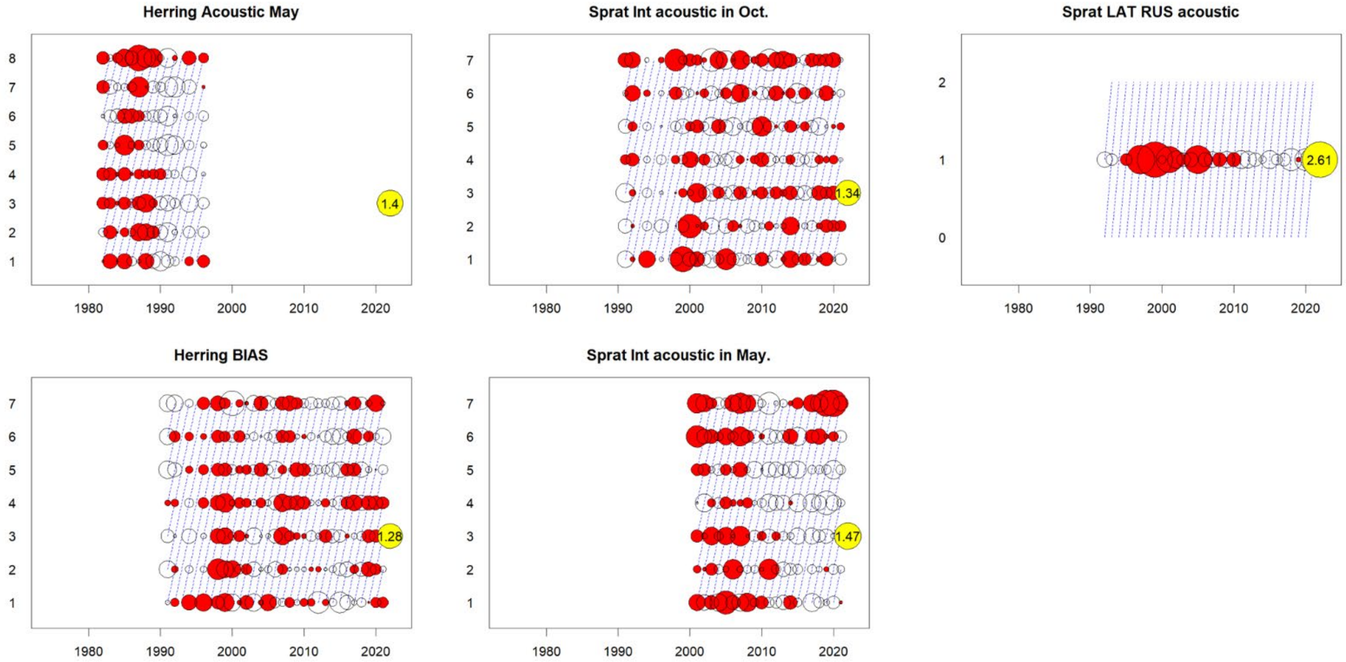


**Figure 6. Survey observations Residuals, log(Survey observed CPUE) - log(expected CPUE). Red is positive, white is negative. The yellow circle illustrates the largest value for reference.**

The fit to stomach data reveals residual patterns, where there is a consistent over- or underestimation (cluster of residuals switching from positive to negative and vice-versa) of other food, sprat and herring in certain time periods (Figure 7). The box plots of residuals also show a clear switch towards positive residuals in the 1990s for cod eating herring, and a clear declining trend for cod eating other (Figure 8). This suggest possible (non-constant) trends in processes such as prey vulnerabilities or availability of other food, which are so far not captured by the model.

**Figure 7. Cod stomach observation residuals (raw = (observed proportion) - (expected proportion)). For residual plots, white is positive, green is negative.**

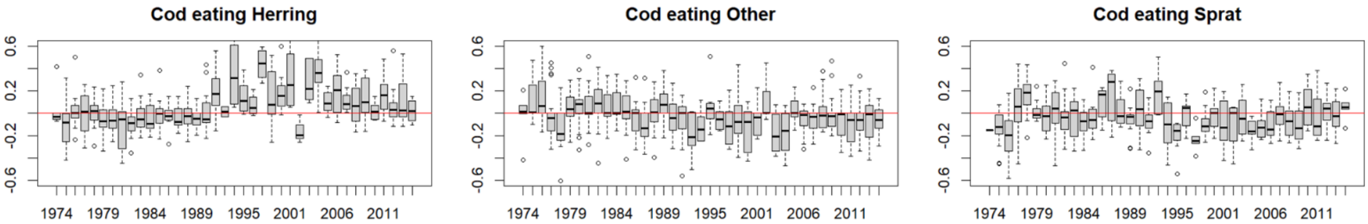


**Figure 8. Box plots of cod stomach observation residuals by year.**

## 1.1.5 Uncertainty

The uncertainty of any one selected output variable in SMS is estimated using the Hessian matrix and the Delta method approximation (see Annex 2). However, we note that these do not include possible error in the input variables, and are thus likely an underestimation of the uncertainties in the model estimates. For example, the stock numbers for cod (a so-called ‘other predator’ within the model) are assumed to be known without errors while considerable uncertainty is associated to the stock assessment estimates. The uncertainty estimates of the model are for this reason considered uninformative and are not presented in detail in the stock annex (see Annex 2) or used further.

Predation mortality (M2) decreases in general by prey size (age) as expected (Figure 9); however for sprat, M (M2 + constant M1) is estimated to be highest for age 1, as was the case in the last key run. This pattern might be explained by a lower spatial overlap between the very small sprat and cod, or may be an effect of the rather uncertain mean weight at age applied for the 0-group.

WGBFAS assumes that mean weight at age in the sea is the same as mean weight at age in the catch. This assumption could be violated for the age 0 individuals, as the representation of age 0 individuals in the catches and surveys are likely the largest individuals in their cohorts. The age 0 individuals are likely underrepresented in the catch and survey data. While the M values needed for the single species assessments start with age 1 and M values for age 0 are less critical, shifts in the cod diet portion of age 0 herring and sprat could lead to changes in M estimates for the other prey age classes. It is therefore recommended to test the robustness of the model outcomes to the values chosen for the mean weights of the age 0-group.

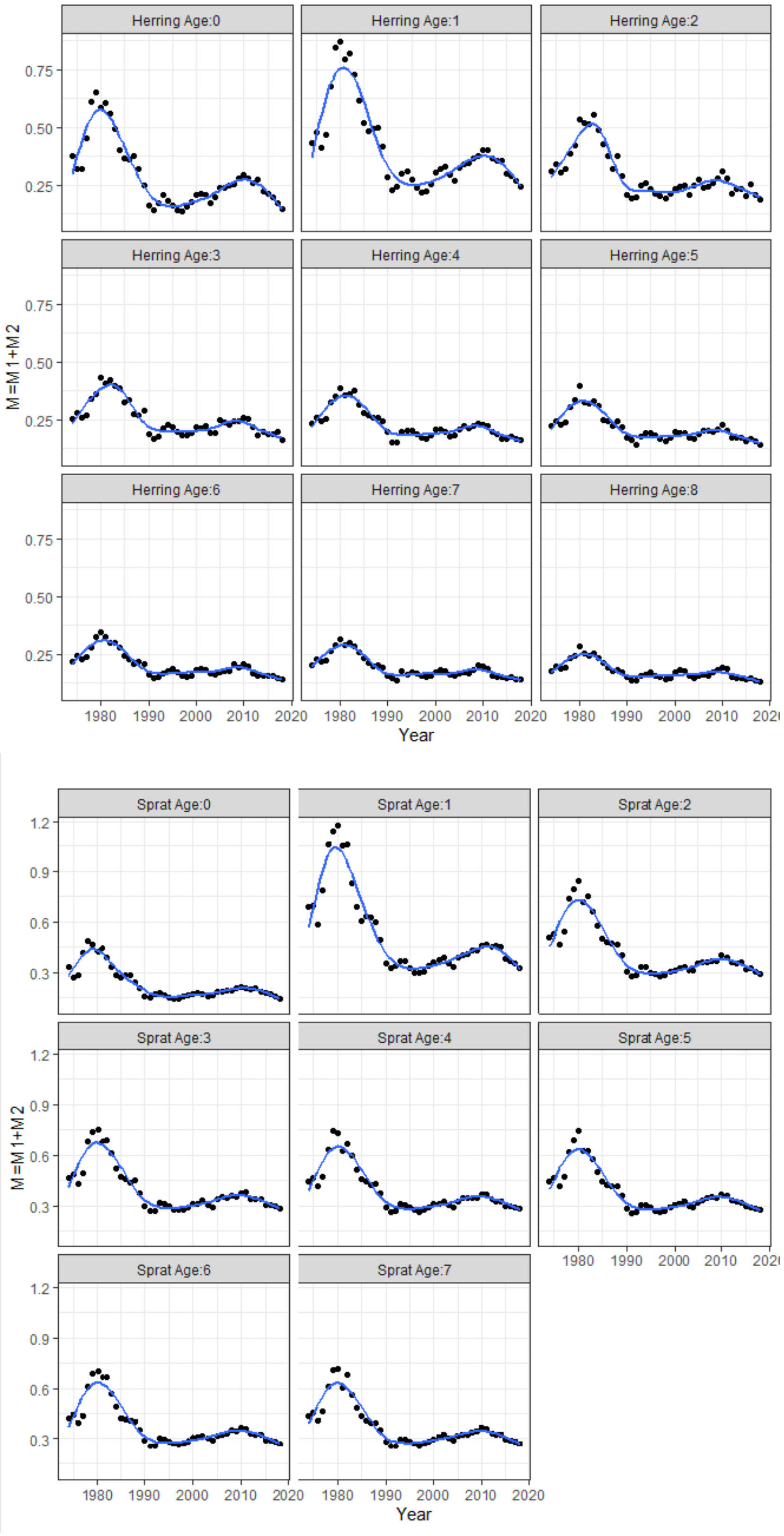


**Figure 9. Annual natural mortalities (M=M1+M2) by species and age. Black dots are the sum of quarterly M1 and M2; the blue line is a gam spline estimate.**

Some of the sensitivities of the model were quantified by running the following protocol:

1. Retrospective analysis (successive 5-year peel of all input data);
2. Robustness test of results towards a change in the residual natural mortality (M1).

In addition, the effects of small model changes that had to be made, such as a new compilation method of the cod stomach contents, a correction of the food ration calculations, and a change in the fishing parameter configuration were tested on the model outcomes, effects of these changes on the model outcomes are described in section 1.1.6. In 2019, several other robustness tests were conducted (Table 1). In the table we present an overview of the tests that were conducted. These were not repeated for the current model input data and not reviewed by the 2022 review team and are therefore not formally part of this review.

The retrospective analysis results in small shifts of F, recruitment and SSB (Figure 10), yet consistent estimates of successive M2 values across clupeid age classes were obtained as additional years of assessment data were added to the key-run configuration (Figure 11). We therefore conclude that these point estimates might be regarded as robust to the variability in assessment data of herring and sprat and not strongly dependent on one specific year of data.

**Table 1. Changes of parameters values tested on the predictions of SMS in 2019 and 2022. The tests conducted in 2019 were not repeated for the current model input data and not reviewed by the 2022 review team and are therefore not formally part of the review.**

| Test | Reference | Tests performed | Executed | Remarks |
|---|---|---|---|---|
| Retrospective analysis | Annex 2, section 5.2.1 | Peeling off 5 years of data | 2022 | |
| Change in M1 value | Annex 2, section 5.2.7 | Half of original value for sprat and herring | 2022 | |
| Change in aggregation period stomach data | Annex 2, section 5.2.4 | 5 or 10 year aggregation period | 2019 | |
| Change in consumption rates | Annex 2, section 5.2.6 | Value at half of original value | 2019 | |
| Overlap index other food | Annex 2, section 5.2.5 | With and without index | 2019 | |
| Change in stomach data | Annex 2, section 5.2.2 | New and old stomach data | 2019 | New data used since 2019 |
| Change in shape size selection function | Annex 2, section 5.2.3 | Uniform, constraint uniform,<br>Log-normal | 2019 | Log-normal size selection was chosen and used in 2022 |
| Comparison with Baltic Sea Gadget mode | WGSAM 2019 | | 2019 | |

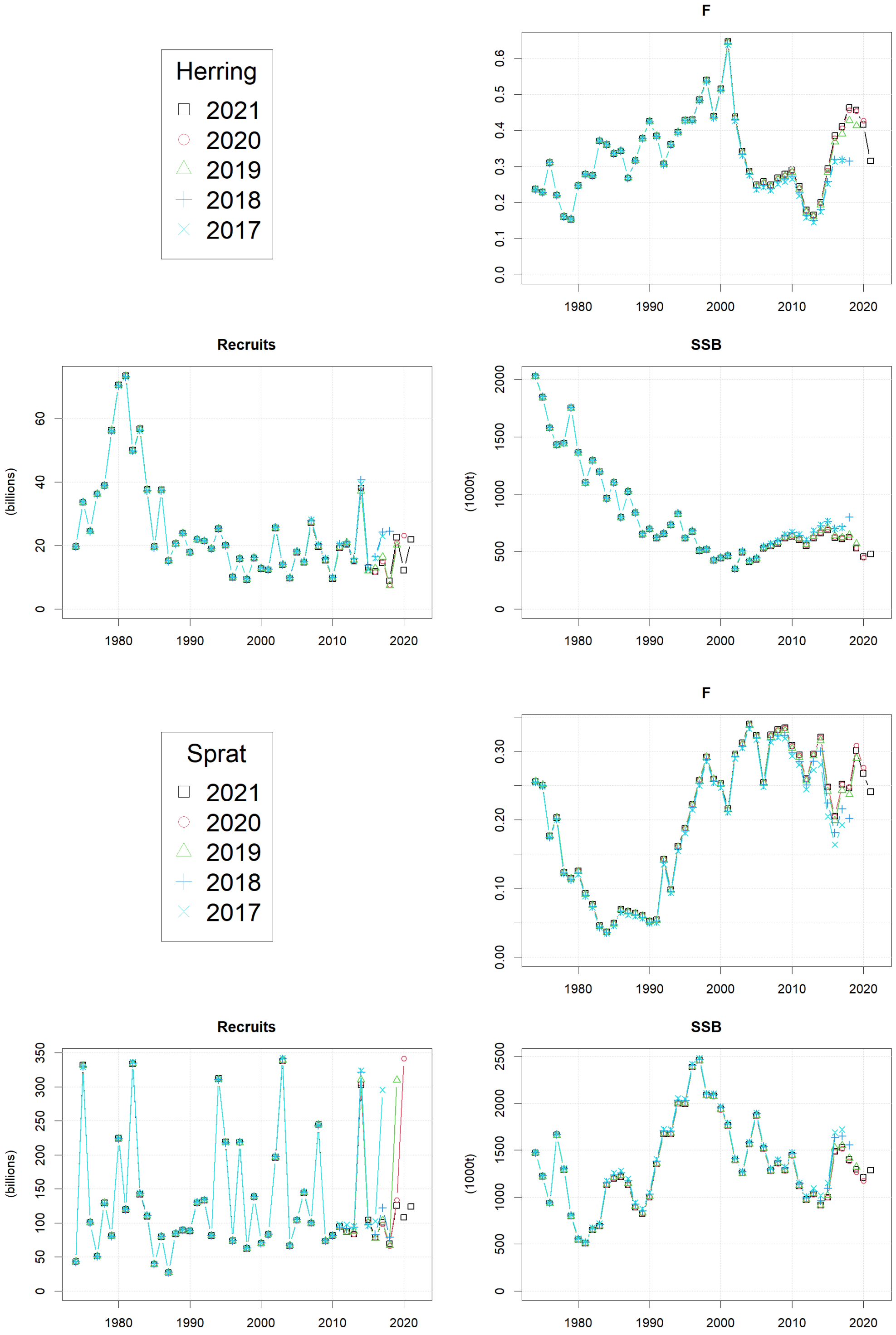


**Figure 10. Retrospective analysis for herring and sprat. Summary output.**

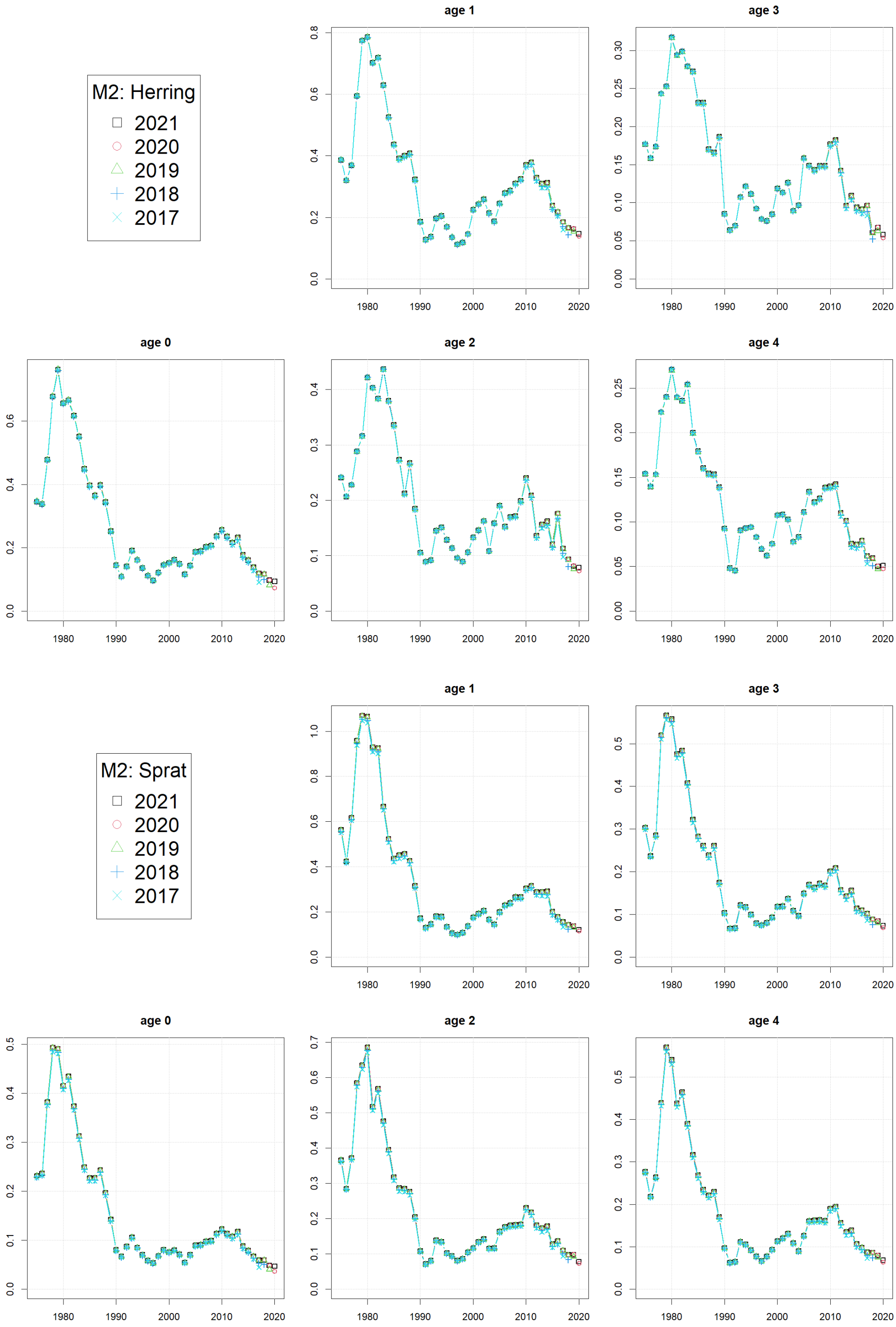


**Figure 11. Retrospective analysis for herring and sprat, M2 at age.**

Residual mortalities M1 are assumed constant and independent of age class in the model parameterization. M1 values for Herring and Sprat are respectively set to 0.1 and 0.2 in this as in the previous key-run; there is no empirical evidence available to support a choice in this value. To test the effect of M1, a simulation was performed with a 50% reduced value for M1 for sprat and herring. The results show a worse fit and higher M2 estimates for both clupeids across all age classes (Annex 2, section 5.2.7, Figure 13). Furthermore, the lower M1 values resulted in lower recruitment, a higher F and a lower SSB estimate for both herring and sprat (Figure 12); though for the latter species this effect is greater due to the larger absolute reduction in M1.

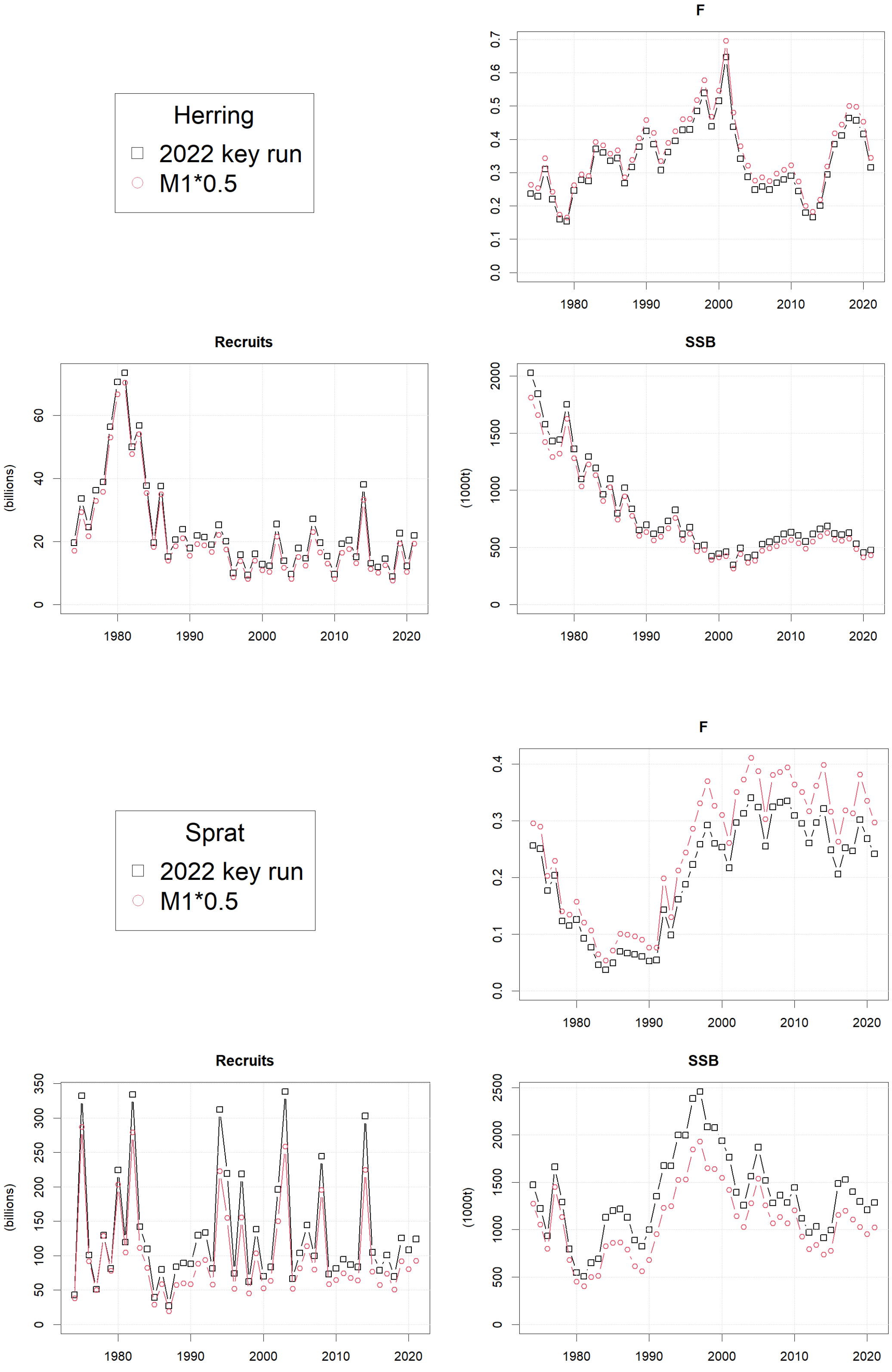


**Figure 12. Effect of reduced M1 values on F, Recruitment and SSB for herring and sprat.**

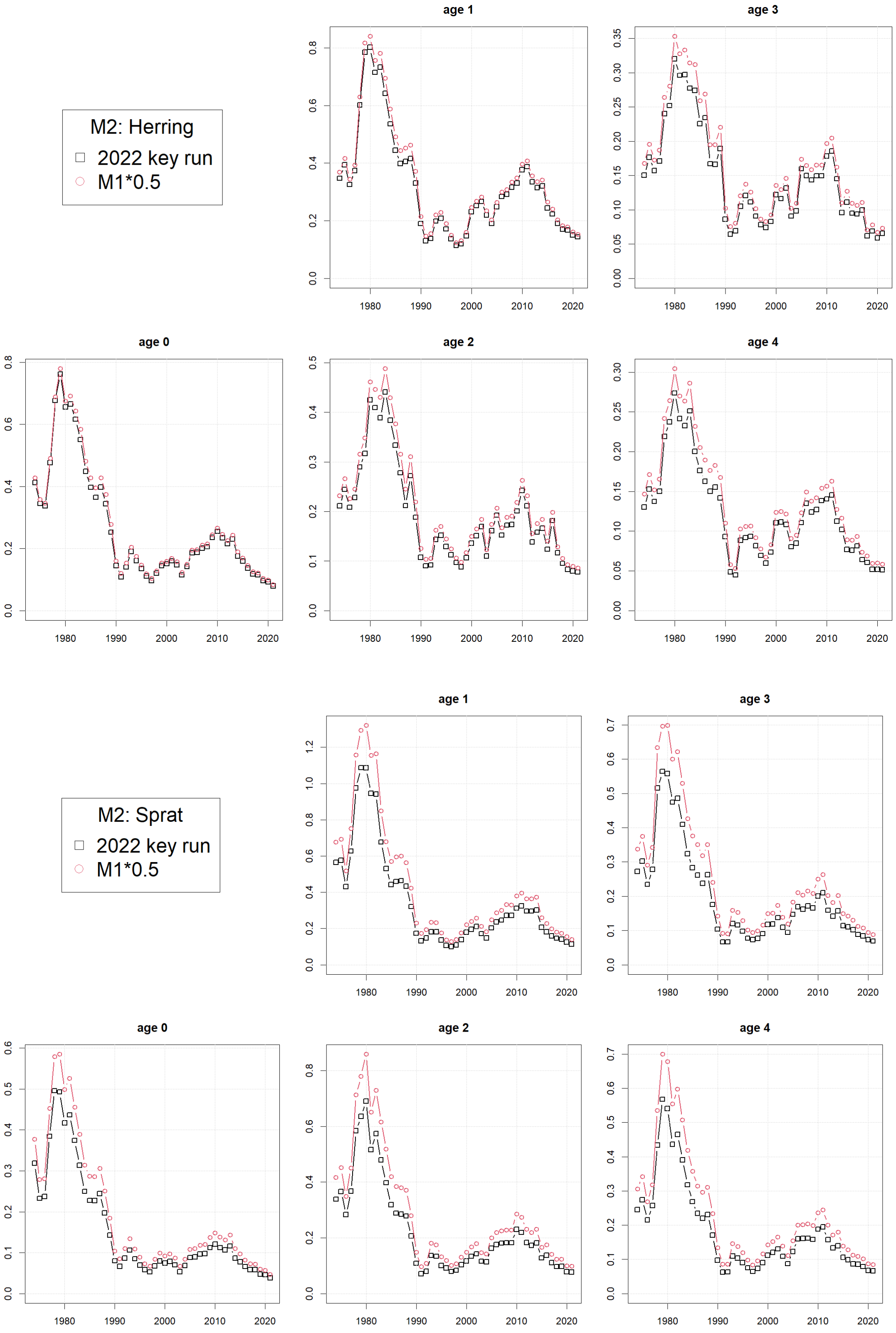


**Figure 13. Effect of reduced M1 values on age-specific M2 estimates for herring and sprat.**

The 2019 review states that the predator-prey size selection has quite a strong effect on old M2 estimates (ICES WGSAM, 2019). In general, the diet partitioning of cod between herring, sprat and the other food component in the model is uncertain. The residuals of the diet data fitting show a significant deviation in the model fit for recent years (section 1.1.4, Figures 7 and 8), which is an indication that the model is perhaps not capable of catching changes in cod diet, such as a shift from benthic to fish or from herring to sprat. Additional analysis of the stomach data may allow for a better understanding of potential changes in the diet of cod and the changes in the model setup that may be needed to simulate these dynamics.

One of the main uncertainties in the model are probably the Cod stock sizes. Due to the issues with the age readings in Baltic cod, estimates for weight and catch at age that SMS needs for its calculations are currently not available. Therefore, cod is at the moment considered as an 'external predator' in SMS. In practice this means, that the cod numbers are kept fixed and assumed to be known without error. In reality, the estimates of the cod numbers are likely to be quite uncertain, especially with the current problems with the assessment of the cod stock in the Baltic (ICES 2022). The effect of different values for the cod abundance on the model estimates should be tested in the future.

### 1.1.6 Previous peer review

The SMS methodology has been reviewed in ICES (ICES 2012a, 2015, 2019).

Compared with the 2019 key-run, the 2022 key-run is based on a revised food rationing, changes in the model configuration for fishing mortality and includes updated assessment data from WGBFAS. Despite these differences, the two key-runs show highly consistent results for the summary output of recruitment, SSB and mean F (Figure 14). For herring, estimates of F, SSB and recruitment are very similar prior to 2016. After 2016, the two key-runs deviate slightly, with the 2022 key-run estimating higher F and lower SSB compared to the 2019 key-run. For sprat, the estimates of F, SSB and recruitment are similar for the entire time-series between the two key-runs, with the 2022 key-run estimating slightly higher F across most years (Figure 14). These differences are likely due to the changes in the fishing mortality configuration settings in the 2022 key-run (Figure 16).

In terms of M2 (Figure 15), the two key-runs show similar results with some discrepancies for certain age groups. For herring, the 2022 key-run estimates higher M2 for age 0 herring prior to 2000 and higher M2 for age 1 herring after 2000 compared to the 2019 key-run. For sprat, the 2022 key-run estimates higher M2 for age 0 herring prior to 2000 (Figure 15). These patterns are likely due to the updated food rationing, where the average quarterly food consumption has been split each quarter according to the distribution key specified by Neuenfeldt *et al.* (2020), see for details below (Figure 17).

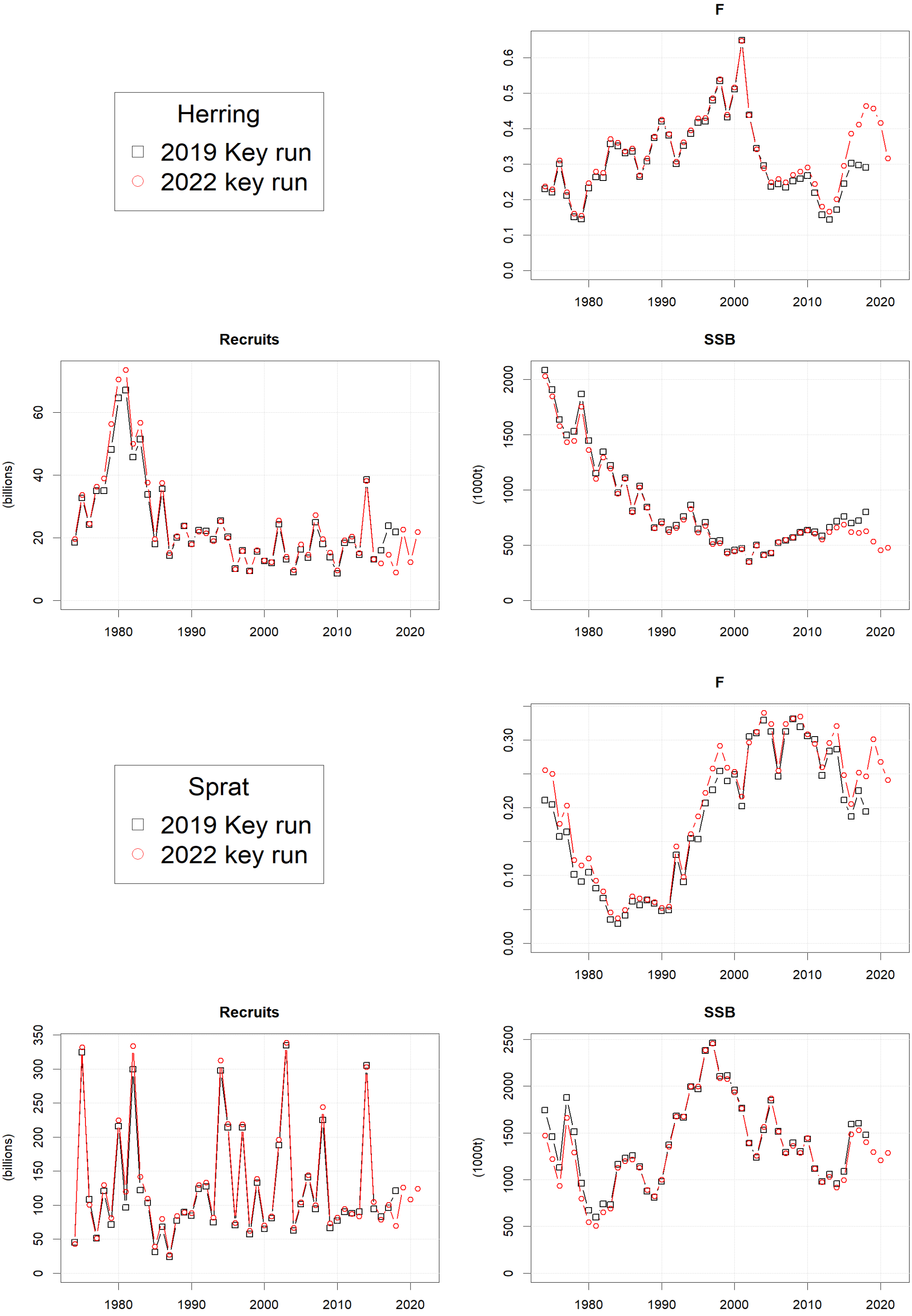


**Figure 14. Comparison of fishing mortality (F), recruitment (Recruits) and spawning stock biomass (SSB) estimates between the 2019 (black squares) and the 2022 (red circles) Baltic Sea key-runs for herring (above) and sprat (below).**

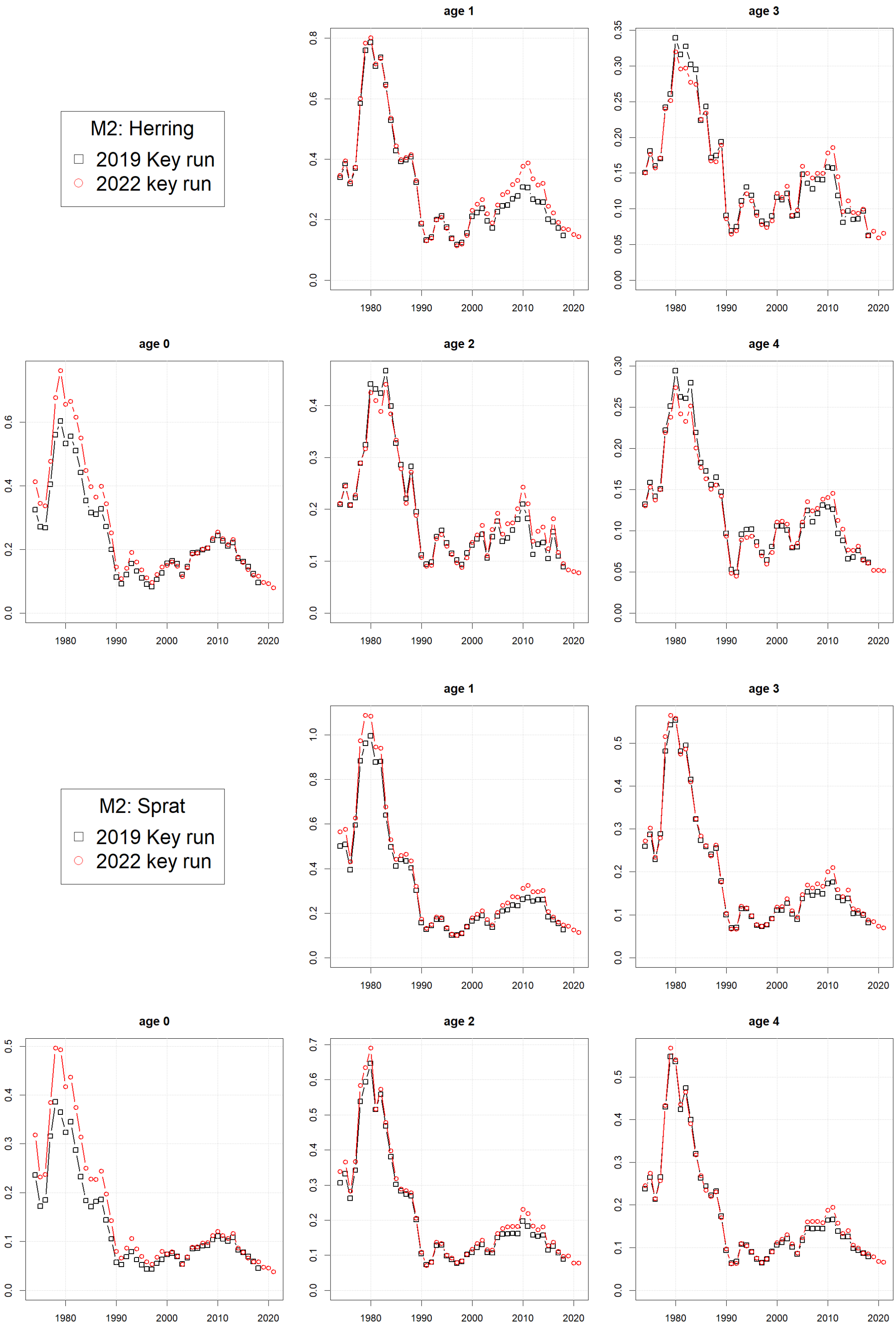


**Figure 15. Comparison of natural mortality (M2) estimates between the 2019 (black squares) and the 2022 (red circles) Baltic Sea key-runs for herring (above) and sprat (below).**

Changes in model parameterization and configuration

For the 2022 key-run, the compilation procedure of the stomach data was performed using the 'FishStomach' R package (https://github.com/MortenVinther/FishStomachs). This introduced some differences in terms of relative stomach contents compared to the beta version of FishStomach used for the 2019 key-run; which, in turn, resulted in negligible deviations of M2 estimates (Figures 16 and 17). The discrepancies between 2019 *vs.* 2022 relative stomach content are likely arising from the allocation of partly identified preys within the latest FishStomach's algorithm.

The 2019 key-run food ration protocol was revised and average quarterly consumption rates were multiplied by 4 to yield average yearly consumption to subsequently redistribute over quarters according to KEY -xyz- in Annex 2. This update resulted in the largest changes observed in M2 (Figure 17). Herring M2 was particularly sensitive to this update across age classes in the recent past (with the exception of perhaps age = 0 Herring from 1974 to 1990). In contrast, Sprat M2 tended to be more sensitive in earlier years of the time-series.

The responses of M2 estimates to the changes in fishing mortality configuration were modest (Figures 16 and 17). For Herring, an extra period was added in the season effect of the fisheries selectivity scenario (*i.e.* 2006–2021) to better capture potential temporal trends in age-specific fisheries effort, and consequently fishing mortalities, over the years. For both herring and sprat, alternative age groupings of the variance on F were chosen to prevent overfitting (see Annex 2 for details). In addition, the catch observation variance was changed from annual to seasonal for sprat. Although marginal differences were found in M2 estimates with this new configuration (Figure 17), the data favours the new fishing mortality configuration (following the rescaling method in Burnham and Anderson (2004) to interpret AIC for model selection); even at the expense of an extra 25 parameters.

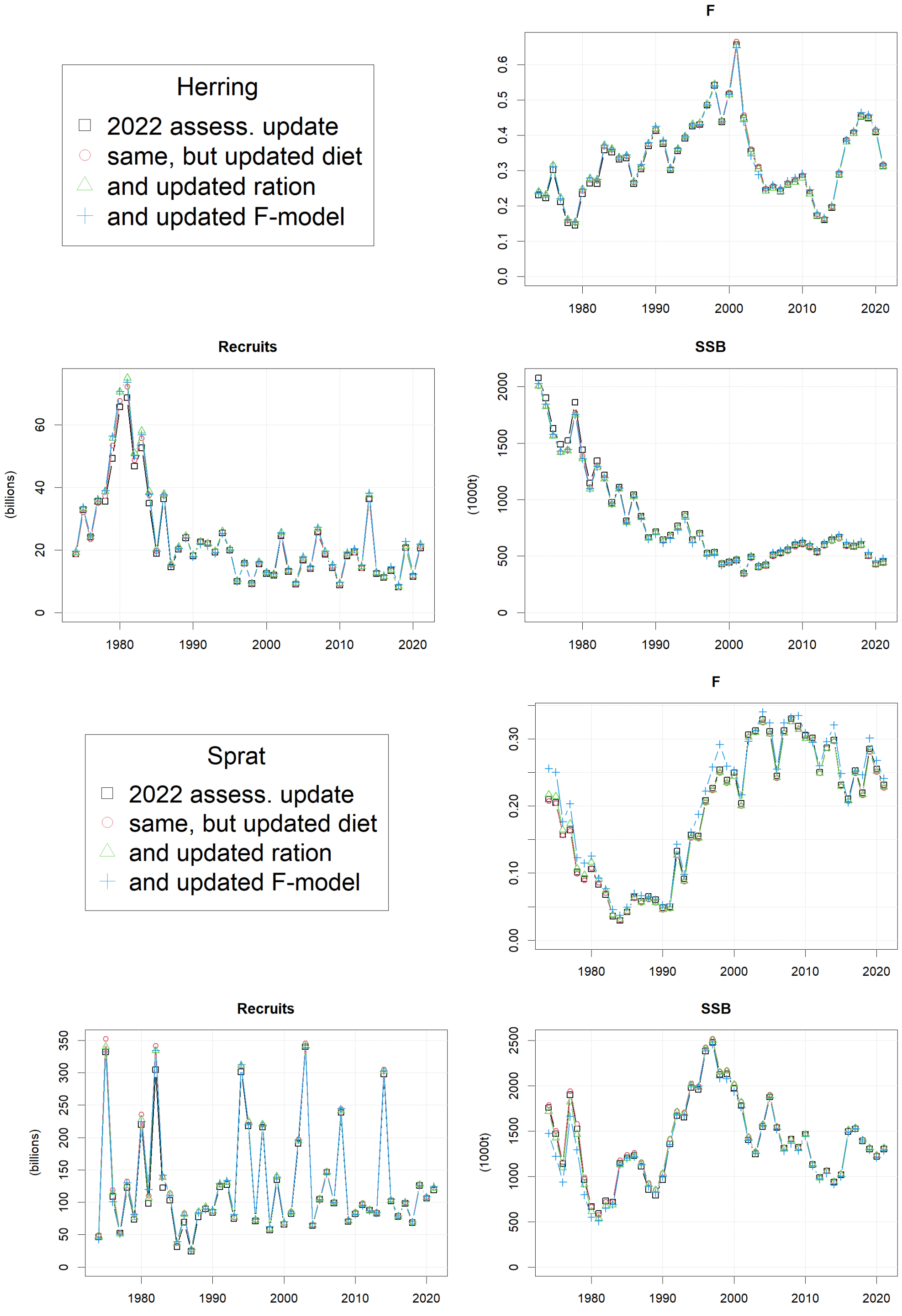


**Figure 16. Comparison of F, Recruitment and SSB estimates for four different versions of the model: 1. With an update of the model input data, 2. With an update diet data compilation method, 3. with an updated ration calculation method, and 4. with a different configuration of the F-model.**

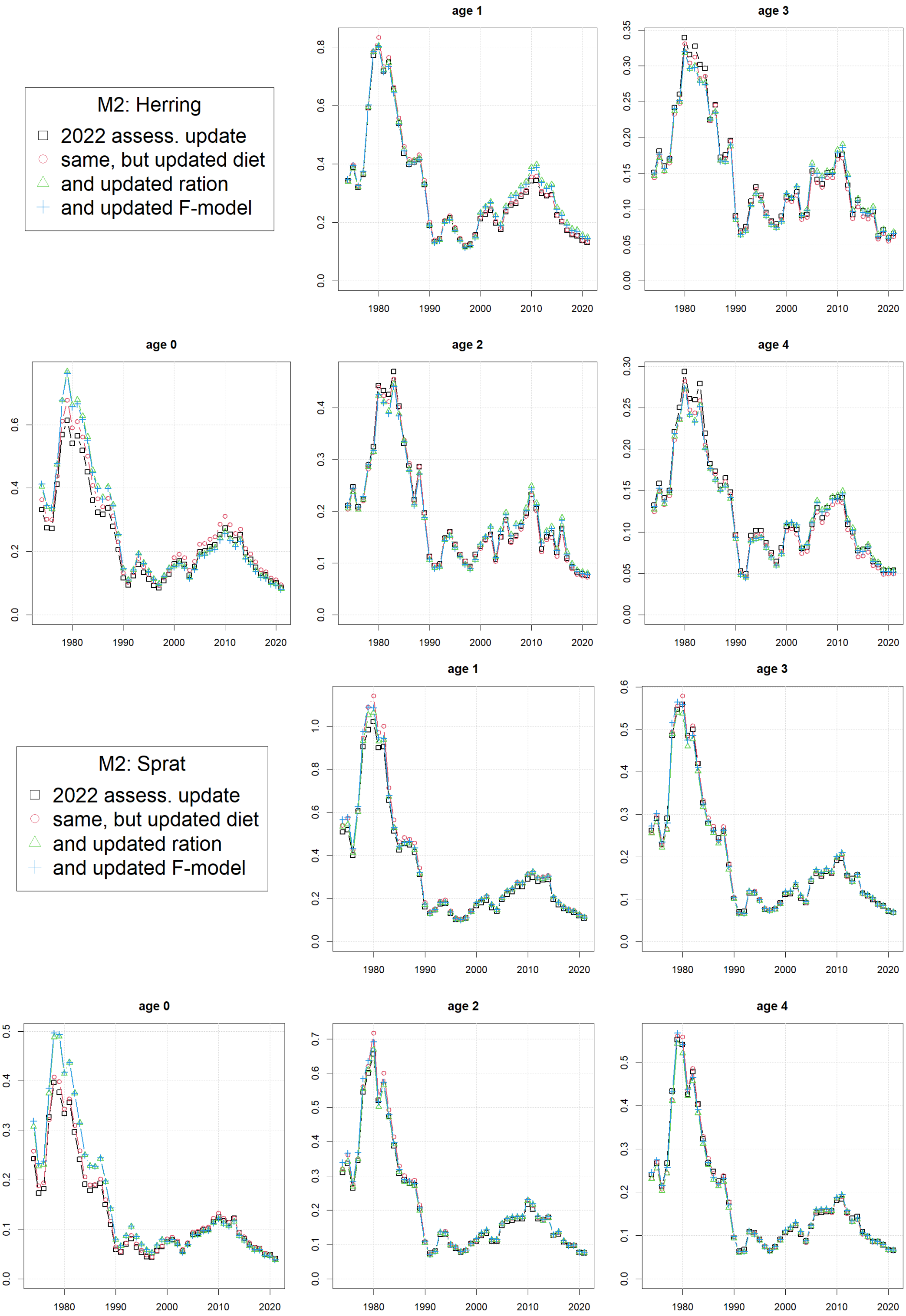


**Figure 17. Comparison of M2 estimates for four different versions of the model: 1. With an update of the model input data, 2. With an update diet data compilation method, 3. with an updated ration calculation method, and 4. with a different configuration of the F-model.**

## 1.2 Review recommendations

WGSAM accepts the model output from SMS as key-run with the settings given in the Stock Annex (Annex 2 of this report).

### Key-run summary sheet

| AREA | BALTIC SEA |
|---|---|
| Model name | SMS |
| Type of model | Age–length structured statistical estimation model |
| Run year | 2022 |
| Predatory species | Assessed species: Herring , Sprat |
| Prey species | Herring, Sprat |
| Time range | 1974–2021. |
| Time step | Quarterly |
| Area structure | Eastern Baltic Sea, ICES sub-divisions 25–29, excl Gulf of Riga |
| Stomach data | Cod: 1974-2014 |
| Purpose of key-run | Making historic data on natural mortality available and multispecies dynamics |
| Model changes since last key-run | All time-series updated. More stomach data included. Stomach data compiled with the FishStomach R package. Cod is now an external predator estimated by WGBFAS Stock-synthesis model. Daily food ration changed for the predator cod. Calculation of rations updated to follow the stock annex (Annex 2). |
| Output available at | Sharepoint/data/Eastern Baltic SMS key-run and https://github.com/ices-eg/wg_WGSAM |
| Further details in | Report of the Working Group on Multispecies Assessment Methods 2022 |

WGSAM considers the key-run as currently best possible run with SMS to provide natural mortality estimates. WGSAM recommends to use these values as input to single species stock assessments. Whether the raw or smoothed natural mortality values should be used in the herring and sprat assessment models is up to the stock assessors and is usually a decision that is reviewed and agreed upon during Benchmarks. The full time-series should be used and not only an update for the years after the last key-run in 2012.

However, there are also clear limitations with the approach and results have been shown to be sensitive to e.g., consumption rates and treatment of “Other Food” as well as the size selectivity of cod (ICES WGSAM, 2019). In addition, the results are likely to depend to a large extent on the outcome of the ICES Eastern Baltic cod assessment. Any bias in this assessment may directly influence the predation mortality estimates. Assumptions around other food and constant vulnerabilities and uncertainties in the body sizes and abundance of age 0 prey individuals may also bias the natural mortality estimates to some extent.

Multispecies models can provide multispecies $F_{MSY}$ estimates to be considered in comparison with single species $F_{MSY}$ values. It is valuable to ask how, for example, the chosen fisheries mortality on sprat and herring may affect the optimal exploitation rate for cod, and vice versa. Multispecies $F_{MSY}$ estimates could be obtained, for example, by concurrently maximising the yield of all species in a system in the sense of a Nash equilibrium, with respect to F. In this Nash equilibrium all ICES single-species MSY reference points can be simultaneously attained and are identical to the multispecies $F_{MSY}$ values. WGSAM can provide the computational machinery necessary to compute such MS-$F_{MSY}$ estimates from any of the key-runs reviewed by WGSAM so far. A graphical representation of a set of $F_{MSY}$ values of all the key runs of multispecies models of the Baltic Sea would (i) facilitate visualising yield trade-offs amongst target species when evaluating other socio-economic objectives and (ii) assist in identifying SS-$F_{MSY}$ reference points that can simultaneously be achieved.

For further work on the Baltic Sea SMS keyrun, WGSAM recommends the following:

1. Additional analyses of the cod diet to get a better process understanding of what is driving the systematic changes in relative stomach contents.

2. A split of Other Food in parts where the time dynamic can be considered (e.g., flounder and *Saduria entomon*) directly or via a proxy and a part that still needs to be assumed constant in time may be beneficial. The availability of the different food groups for cod in the Baltic Sea have changed strongly over time (e.g. Haase *et al.* 2020).

In 2019, the WGSAM reviewers made these two recommendations, on which no work has been done since then. The overestimations of the model in the amount of herring eaten by cod and the underestimations of other food and sprat (all in recent years, Figure 8), indicate that it is still a good idea to consider these recommendations. The recent changes in the condition of Baltic cod indicate that there has been likely a change in the food availability for cod. Although it could be difficult to implement these changes they could be quite important for reliable estimates of the model.

3. The inclusion of spatial dynamics in the food consumption of cod (either directly or via overlap coefficients) may improve the fit to data sources.

In 2019, the WGSAM reviewers made this recommendation, which has not been considered so far. As long as the spatial overlap between prey and predator is constant through time, incorporating an overlap coefficient should not have a strong effect on the model predictions as this is currently addressed via the vulnerability parameter. However, if there is a temporal trend in the spatial overlap between predator and the prey species, and/or if the prey age classes are unevenly distributed through space, an overlap coefficient (per prey age class) could make a difference for the model predictions.

4. A run with age 1 as recruits could be tried because input for the 0 group is highly uncertain.

In 2019, the WGSAM reviewers made this recommendation, which has not been considered so far. The idea is that the weight at age data for the age 0 classes of the prey species are so uncertain that the model cannot estimate the consumption correctly based on the data that is currently used. However, this may be a difficult change for the model as age 0 herring and sprat form an important part of the cod diet. It is relevant to highlight the uncertainty and potential issues with the quality of the estimates for the age 0 prey classes before they could be used as input in a single species stock assessment. Discussion with the stocks experts from the assessment working group on possible alternative data sources on the weight-at-age of age0 clupeids would be beneficial.

5. Account for the uncertainty in cod number at age in the model.

In 2019, the WGSAM reviewers made this recommendation, which has not been considered so far. Due to the current problems regarding the age-reading of Baltic cod, an age-length-based model is used for the stock assessment. Cod number-at-age in the stock are retrieved from the stock assessment and used to represent the cod stock in the SMS keyrun without uncertainty.

6. Provide a multispecies $F_{MSY}$ estimation that can be used as a comparison with the $F_{MSY}$ values from the stock-by-stock (SS) frameworks currently used for advice.

The current model version of SMS of the Baltic Sea is not suitable for predictions of multispecies $F_{MSY}$ values. Unfortunately, due to the non-dynamic nature of cod in the 2022 key run, it is not possible to predict how fisheries mortality may affect the other species in the food web through indirect effects. It is advisable to calculate this metric in the future if possible. The consideration of optimal yield trade-offs amongst the commercially exploited species in the complex is an important facet in improving fisheries management advice and an ecosystem-based approach to fisheries management.

## 1.3 References

Andersen NG, Beyer JE (2005a) Mechanistic modelling of gastric evacuation applying the square root model to describe surface-dependent evacuation in predatory gadoids. J Fish Biol 67: 1392−1412

Andersen NG, Beyer JE (2005b) Gastric evacuation of mixed stomach contents in predatory gadoids: an expanded application of the square root model to estimate food rations. J Fish Biol 67: 1413−1433

Haase, K., A. Orio, J. Pawlak, M. Pachur, and M. Casini. 2020. Diet of dominant demersal fish species in the Baltic Sea: Is flounder stealing benthic food from cod? Marine Ecology Progress Series 645.

ICES 2012a. Report of the Workshop on Integrated/Multispecies Advice for Baltic Fisheries (WKMULTBAL). ICES CM 2012/ACOM:40

ICES 2012b. Report of the Working Group on Multispecies Assessment Methods (WGSAM). ICES CM 2012/SSGSUE:10

ICES 2015. Report of the Working Group on Multispecies Assessment Methods (WGSAM). ICES CM 2015/SSGSUE:10

ICES. 2019. Working Group on Multispecies Assessment Methods (WGSAM). ICES Scientific Reports. 1:91. 320 pp. http://doi.org/10.17895/ices.pub.5758

ICES 2022. Baltic Fisheries Assessment Working Group (WGBFAS). ICES Scientific Reports. 4:44. 659 pp. http://doi.org/10.17895/ices.pub.19793014

Lewy, P., and Vinther, M. 2004. A stochastic age-length-structured multi-species model applied to North Sea stocks. 20. 33 pp.

Neuenfeldt, S., V. Bartolino, A. Orio, K. H. Andersen, N. G. Andersen, S. Niiranen, U. Bergström, D. Ustups, N. Kulatska, and M. Casini. 2020. Feeding and growth of Atlantic cod (Gadus morhua L.) in the eastern Baltic Sea under environmental change. ICES Journal of Marine Science 77:624–632.

# 2 Georges Bank review

This review concerns the evaluation of three modelling frameworks presented during the WGSAM which took place in Woods Hole on 10–14 October 2022. The three models are presented as a suite of tools to address long-term multiple management objectives in the context of ecosystem-based fisheries management for the Georges Bank.

The three models consist of a multispecies production model (MSSPM), a multispecies length-based model (Hydra), and a mass-balanced ecosystem model (Rpath). Overall, the aim of Hydra and MSSPM is multi-species assessment and catch advice for a subset of focal species by taking advantage of the structural difference of these two models. Differently, the overall aim of Rpath is the quantification of the foodweb interactions at large and of whole ecosystem consequences of alternative management strategies.

It is important to remark that for WGSAM, multispecies model keyruns remain complete model implementations used to support the advice process. While none of the three models reviewed here is at that stage to be evaluated as a keyrun, the review detailed in the following sections is largely based on the review criteria (at least those applicable given the current status of the model) adopted by WGSAM for evaluation of model keyruns (ICES 2019 and also found at https://ices-eg.github.io/wg_WGSAM/ReviewCriteria.html).

This review starts with general considerations common to each of the three models regarding the spatial domain and the centralized treatment of the data before detailing on each specific model.

### The spatial footprint and place-based management

Georges Bank is considered an ecosystem production unit (EPU) based on a combination of bathymetry, bottom sediments, temperature, salinity, and primary production characteristics. The relatively shallow waters and hydrographic characteristics make Georges Bank a hotspot of productivity in the Northeastern Atlantic shelf which supports the idea for placed-based management and for the development of fisheries ecosystem plans specifically for this area.

The boundaries of Georges Bank, as defined by the NEFSC bottom trawl survey strata, are used as a common spatial footprint for the implementation of a multispecies surplus production model (MSPM), a length-based multispecies model (Hydra) and a foodweb model (Rpath) here evaluated by a subgroup of experts within WGSAM. It is important to note that this review does not concern the strength and weakness or appropriateness of the Georges Bank EPU for the development of fisheries ecosystem plans but explores/evaluates the implications and challenges that this may represent for the implementation of the three models proposed for this keyrun.

The conceptual models for all the three models implemented assume Georges Bank as a closed system which might be problematic because the definition of its boundaries as an EPU is limited to considerations on the primary productivity of the system and neglects that secondary production is likely to be influenced by contribution from areas outside the EPU. The distribution of piscivorous fish species shows a clear increase in abundance towards the boundaries of the study area suggesting potential issues of spatial containment when modelling trophic interactions especially in the upper part of the foodweb. Of the 8 teleosts and 2 elasmobranch species included in the MSPM and Hydra models only the stock management unit of cod and two flounders show a good correspondence with the EPU. All the other stocks show considerably larger distributions and migratory species like herring and mackerel are known to be transient within the EPU. The mismatch between the GB and the stock-area definition is well reported but it would be useful to evaluate how the life cycle, dynamics and productivity of each stock (assumed to be a biological unit) relate to the EPU and to the main assumptions of the models. Leading questions for further development of the work include:

- what proportion of the catch and fishing mortality occurs outside the EPU?
- what proportion of natural mortality occurs outside the EPU?
- what proportion of the energy intake is provided by preys outside the EPU?
- do these proportions change over time?
- what is the population structure (age/size) of those stocks inside and outside the EPU and does it suggest some level of separation for spawning components in the GB?

Material to start answering these questions has been already presented during the WGSAM meeting indicating that catches taken on GB may represent <50% of the catches, hence only a limited part of the total fishing mortality for some of the key stocks analyzed by the models. Moreover, several of the species considered occur in Georges Bank with two or more stocks. Where information is available, we recommend attempting to answer the leading questions above approximating the stock level.

A possible conclusion from answering these questions could be that from an ecological and modelling perspective it would be difficult to understand the dynamics of the higher trophic levels in GB outside the context of the Northeast continental shelf. In that case, expanding the models outside the boundaries of the EPU, and/or explicitly accounting for the input/output of fish and energy across the boundaries will likely be needed.

Species selected: MSSPM and Hydra used a common set of 10 species. They were selected primarily because of their historical commercial importance and for management reasons. It would be useful to add ecological considerations to evaluate if other species should be included or given priority for inclusion in the models. That would include both predators as well as preys currently part of other food which includes a number of assessed and managed species.

### Data

Dedicated R packages have been developed to handle both commercial data (comlandr), survey data (survdat), and compositional information (mscatch) used by the three models. This improves considerably reproducibility and speed of the workflow as well as future updates of the dataset and models.

### Diet data

Stomach data are available from different seasons with the majority of stomachs collected during Spring and Fall surveys. The underlying data and to the extent possible assumptions on trophic interactions should be consistent among the models (from the objective 2) to enhance comparison of results. Given that the production model does not fit to diet data, it might be difficult, however, to get a converged run if the interactions are complex.

The interactions need to be double-checked because some species are considered predators only because of occasional occurrences of the other species in their stomachs while they are mainly prey (e.g., flounders). It is worth reducing the number of trophic interactions to those that really matter for the prey dynamics and fitting. It is difficult to come up with an automatic method for choosing the trophic links between the stocks. With other models in WGSAM, this is usually done by visual investigation of the diet.

A lot of the other food is actually data we have information on at the family or group level. This is considered other food in the two estimation models while a different approach is taken with Rpath that can include families or functional groups. Ideally, it would be good to make some assumptions on how to allocate these proportions to utilize this information in the estimation models as well and aim for the highest possible level of consistency in the treatment of the stomach data across the three models used

**Catch data**

A great amount of work has been produced in order to extract catches that occurred on GB from catch data which are reported at the scale of statistical areas and that do not match the borders of the Georges Bank spatial footprint. The estimation method, based on the use of VTR and observers data, is sound but the availability of VMS data would help considerably the allocation of catches on the GB. Once time-series with the annual proportion of catches per species and statistical rectangle are calculated their averages are used to split the catch and compile input catch data for the models due to little trust in the VTR data. It is difficult for the review team to judge the quality of the VTR data and the need to take an average proportion for the entire time-series but it would be relevant to consider alternative options like a moving average on the annual proportions and/or applying an average only on the older period of data. Resulting alternative time-series should at least be compared. This is especially important for those stocks where a consistent trend in the proportion of catches within the GB EPU has been observed over the historic period. The whole procedure for splitting catches should also be repeated separately for the different times of the year, gear and fleets at least to check for seasonal variations in the fishing patterns.

Comparison of catches inside and outside the GB from rectangles matching the spatial footprint (521, 522, 525, 526, 537, 538, 551, 552, 561, 562) show that for some species a considerable proportion of the biomass is taken outside the EPU. This poses some questions on the operationality of the method in terms of monitoring compared to defining the area according to the statistical areas.

Evaluation of the catches estimated from the EPU compared to the total catches of the stocks is considered relevant even if the comparison requires merging catches for the multiple stocks of the same species occurring in GB.

## 2.1 MS production model

### 2.1.1 Is the model appropriate for the problem?

MSSPM is intended to provide stock assessment for the 10 modelled species, including catch advice by species. It is also designed to, potentially, quantify predation and competition interactions and estimate carrying capacity by species, trophic guild, as well as for the entire modelled system. This model is not designed for advice purposes in isolation, but to be a part of an ensemble of multispecies models. It has been implemented in an interfaced software, where the estimation procedure can be launched and monitored.

Production models are currently used in situations with sparse dependent and independent fishery data, and shortage of information about the biology of the stock. In the case of the Georges Bank fishing ground, with relative abundance of high-quality data, this is not the main reason for the development of this multispecies production model, but having a relatively simple multispecies model that can provide preliminary estimates of species interdependent dynamics. More importantly, in more complex versions of the model not presented during the meeting

(Gamble & Link, 2009; 2012), a competition parameter by pair of prey-predator species, and a common carrying capacity by trophic guild and a global carrying capacity parameter is also estimated, which can be of relevance for an EBFM in the Georges Bank.

A candidate model fit was not presented due to convergence issues. Hence, the assessment of the model performance and analysis of diagnostics was not possible. However, based on the documentation and the results presented, MSSPM does not seem to be ready to be used in stock assessment, and it may be necessary to modify some elements in the optimization process and the model structure (see sections below).

### 2.1.2 Assumptions (scientific basis, computational infrastructure; adequacy of conceptual model)

The model (and its variants) has been published (Gamble and Link, 2009, 2012), and the equations derive from well-established ecological principles (Schaefer, Lotka-Voltera, Verhulst-Pearl) also implemented in other multi-species models.

Several forms of the model are available in the software, a simple version was favoured as a first attempt to model the GB system, assuming predation and density dependence at the species level (so no competition) and some environmental covariates. The approach of sequential increase in model complexity and fitting (starting by fitting a single species version of the model on each species) is appropriate.

It is worth noting that the assumption of density dependence and environmental effects differ from usual MS assessment models and from the two other models developed for the region.

As a production multispecies model, the dynamics of the modelled populations are driven by the catch, predation interactions, the species growth rate and carrying capacity. Allowing extra flexibility in the input and output of biomass on each population (other than the biomass produced as result of population growth determined by growth rate and carrying capacity) would potentially help in fitting the observed biomass indices as well as accounting for lost and gain of individuals under the problematic assumption for the Georges Bank of closed populations. This flexibility could be achieved by incorporating a random process error in the model (see suggestions in sections below).

**Fitting procedure**

The model estimates values for r (population growth rate), B0 (initial state biomass) and K (carrying capacity) parameters per species and predator prey interaction coefficients (rho). The estimates of the parameters are derived through numerical exploration of the parameter space (metaheuristics). Several algorithms are available and tested, the stopping criteria being a tolerance value. The procedure is sequential, with a first approximation of r, B0 and K estimated in single species models and used to refine the search domain when fitting the MS model.

The first attempts did not show convergence and the experts recommended several actions to improve it and evaluate the performance:

- Carry out a sensitivity analysis of the objective function to the parameters, in order to refine the bounds of the parameter space to explore and identify possible local minima.
- Standardize parameter values that are on a very different scale to improve the chance of convergence
- The modelers used several algorithms to fit the model to data and this can be pursued, but the choice of the algorithm could efficiently be guided by the characteristics of the

optimisation problem and also sequentially combining global and local search algorithms.

- As part of best practices when using metaheuristics, it is recommended to test that the estimation is independent from the settings by repeating the estimation several times with different starting values, boundaries of parameter space and metaparameters (specific parameters of the algorithm such as mutation rate, tolerance, etc.) to assess the robustness of the estimation.
- Given that the convergence is not guaranteed when using such optimisation methods, various diagnostics are recommended to assess the quality of the estimation. Traces of parameter and objective function as well as correlation matrix between parameters are good tools to inform on the quality of the convergence and should be provided.

The modelers were faced with difficulty fitting the model and getting it to converge despite using different metaheuristic algorithms. While it might take a considerable amount of work, it might be useful to try automatic differentiation to fit this model (e.g., autodiff, ADMB or TMB), as it may help the fitting procedure. A source of inspiration could be the SPiCT model currently used in ICES.

The MSSPM aims to add to the NOAA toolbox a simple, fast and user-friendly tool. The graphic user interface is easy and intuitive to use, making the tool accessible also to less experienced modelers. It seems useful to solve the estimation issues of the model before further development of the GUI.

### 2.1.3 Is the input data quality and parameterization sufficient for the problem?

It would be useful to document why certain interactions are estimated over others and possibly harmonize these decisions with the Hydra model to ease comparison between the models.

The objective function is currently the sum of squared errors summed over all species. This is equivalent to having a Gaussian likelihood, with the same variance on all species. It possibly gives more weight to species with high biomasses and a standardization (scaling by the average biomass of the species for instance) would prevent the algorithm from fitting these in priority.

The objective function uses only the fall survey data. It could advantageously be completed with the spring survey data, so the model represents a compromise between the two seasons more easily comparable to Hydra. It has also been discussed whether using catch as observations and using effort data as input could be more informative to the optimisation. The compromise between on one hand, additional information, and on the other hand, additional parameters (catchability), to estimate the quality of effort data (and their standardization across fleet and gear) has to be evaluated. Along the same lines, it should be evaluated whether diet data could be used in the objective function.

### 2.1.4 Comparison with observations

A fitted model was not presented, therefore it was not possible to assess how well the model fit observations.

### 2.1.5 Uncertainty/sensitivity analyses

Sensitivity analysis to the fitting procedure as described above should cover initial values, meta-parameters and stopping criteria.

The sensitivity to the fitting to the data weighting and influence of the different datasets (i.e., leave-one-out test) could be evaluated to better understand model response to the different data components.

If a new fitting framework was to be adopted a comparison of the estimates between the different methods would be of interest.

The uncertainty could be calculated using a Bayesian framework, which would probably involve running a Markov chain Monte Carlo procedure (see for example Gelman *et al.* 2013). However, this may not be feasible and an alternative could be to calculate the Fisher information, the hessian of the log likelihood at the maximum likelihood estimation to get uncertainty in the model parameters. The delta method (Oehlert, 1992) could be used to determine uncertainty in the variables of interest that are not parameters.

## 2.2 Hydra

### 2.2.1 Is the model appropriate for the problem?

The main goal of this review was to investigate if Hydra could be used as an assessment model with the goal to provide catch advice in an Ecosystem-based fishery management (EBFM) context. This review considers therefore the model structure, data availability and use in the model, estimation performance, and treatment of uncertainty and sensitivity analyses.

The model is a length-structured model coded in ADMB. It was initially made as a simulation model and published in Gaichas *et al.* (2017). For this review, the model was extended to allow parameters estimation via negative log likelihoods.

Length-structured models are notably relevant as catch and predation are usually length-based processes. Moreover, length-based models are particularly useful when information on age is limited or uncertain, although in Hydra, transition of fish across length bins is still based on the von Bertalanffy equation that necessitates some information on growth as a function of age.

The model currently uses a limited number of length bins (5 for all stocks). While this number can be easily increased, the model uses less information than an age-structured model (except for the diet information) when the number of bins is less than the number of age classes in the single stock assessment. It seems reasonable to have at least the same number of length bins as the number of age classes, if not more.

Some limitations were raised during the meeting regarding the model assumptions, model structure and data. This is detailed fully in section 2.2. These limitations need to be addressed before Hydra can be used for catch advice in the future.

### 2.2.2 Assumptions (scientific basis, computational infrastructure; adequacy of conceptual model)

**Length bins and adaptive time step**

The use of 5 length bins seems too coarse to capture the complexity in the size structure of most stocks (for instance, can we capture the transition throughout the stock size/age structure of known large year classes?). Moreover, precision in the estimation of the selectivity of a fishery or of the prey-size preference of a predator (at the moment not estimated within the model) can be compromised by a too coarse size resolution in the model.

Considering the differences in the length span of the various species, it would be convenient to be able to set the number of length bins specifically for every species.

At present, the time step of the model is a function of the length bin of the fastest growing species in the model. Such dependency may represent an important limitation for the implementation of models with more resolved size compositions because it would automatically increase the time resolution of the model beyond the resolution of available data and result in a disproportionate increase in the computation time. Ideally, it should be possible to define the time step resolution in the model independently from the length step resolution or at least relax this dependency (see Spence *et al.* 2020). The transition matrix (proportion of fish that change bin at each time step) is an input of the model currently assessed based on von Bertalanffy growth. The growth parameters and the proportion of fish that change length bin at each time step are estimated outside Hydra which limit its ability to include uncertainties in growth.

In addition, heterogeneity in growth is caused by the bin sizes and the time-step of the model. It is worth checking whether the length variability created by the model reflects the population variability. It could be of interest to account for the individual variability around the growth curve.

At present the model does not allow for changes in growth over time. Growth is unlikely to be constant, as a strong decline in length at age has in fact been reported for several stocks in the Georges Bank. This could limit the ability of the model to fit length composition data adequately and most importantly, would result in a discrepancy between the model assumptions and the data. The importance of accounting for changes in growth over time is especially important if the number of size bins is increased in future versions of the model.

### Selectivities and annual catch (important with regard to catch advice)

Fishery selectivity is fleet specific and estimated as a function of size. For each fleet, the model estimates a species-specific catchability parameter but the same size selectivity for all the species caught by the fleet. The shape of the fish, swimming ability and other factors such as the availability of different size groups (i.e., due to ontogenetic changes in species habitat selection) can result in a different selectivity by species for a given fish length. It is expected that this problem will be more important with the increase in the number of size bins (higher size resolution) or time resolution of catch data. It is recommended to estimate a selectivity per stock. An easy way to do this without changing the model would be to assume one fleet per stock, however this would limit the ability of the model to account for technical interactions if the fleets are not correlated in some way.

The information on commercial catches is provided to the model annually despite most Hydra implementations having multiple time steps per year. At present the model assumes that fishing mortality is constant within the year, hence the same F is assumed across the time steps of a year. Deviations from this assumption, as in the case of pronounced seasonal patterns in the fishing effort, might result in biased estimation of F. It is recommended to relax the assumption of a constant F and allow the model to input catch data by time step. Considering the current approach in the definition of the number of time steps, the preparation of input catch data might be more demanding given that new catch data need to be compiled with changes in the time steps. However, the streamlined procedure for the data preparation supported by the comlandr R-package is expected to considerably simplify the task.

### Survey timing

Survey timing in the model should be set to the closest time when the survey occurred instead of being fixed to the middle of the year to account for both mortality and growth of fish within the year. Forcing the time of the survey to the middle of the year is likely to increase the potential conflict among different data sources as in the case of multiple surveys (i.e., the Spring and Fall surveys in the Georges Bank) and in general deteriorate the fitting.

### Predation and fitting to the stomachs (incl. vulnerabilities assumptions, scaling of "other food")

Predation is represented in the model by three main processes: (1) consumption, (2) prey species preference and (3) predator-prey size selection. In the current implementation, the model fits stomach contents of the predators in terms of proportion of the modelled prey and other food in the stomachs, and estimates only parameters in the process (2). Stomach data are currently fitted using a Multinomial distribution which is problematic given that the distribution sums to the total numbers while stomach data are proportions that sum to 1. A Dirichlet distribution or another distribution that supports proportions as observations should be used instead and would provide a better fit to the diet data.

In addition, the parameters for vulnerability of a prey to a predator in the model are assumed known and equal to 1 if the prey is eaten by the predator. The M2 equation implies that the vulnerability for other food is also 1. In multispecies statistical models, the vulnerability parameters are usually estimated and scaled in some way to the other food. This allows the diet proportions to inform the predation equation via the predation parameters, including the vulnerability. Not estimating the vulnerability parameters per prey might leave no room for the model to fit to the stomach content. We suggest to lift this assumption and estimate the vulnerability parameters in the model and make them relative to the other food.

Per capita consumption uses species-specific mean stomach weights taken from Bowman and Michaels (1984). If information on the total stomach weights is available, at least from part of the stomach data and if needed from a broader area on the northeast continental shelf (NEUS), it would be relevant to evaluate/corroborate these old literature values with more recent observations. Food intake parameters are only separated for elasmobranchs and bony fish following recommendations in NEFSC 2010, which is an appropriate initial assumption. However, a model implementation that can be used for applications, such as for instance a future ICES keyrun, would require more details in the intake parameters based on a broader evaluation of the available literature including experimental work on some families like gadoids (i.e. see Jones 1978 as an example on the consumption of cod and haddock).

Predator-prey size selection is a function of the predator-prey weight ratio which is assumed to be equal for all the prey species of a predator and at present it is not estimated within the model. The assumption of a predator-prey weight relationship common for all the preys of a predator may work as a first assumption but it is recommended to evaluate it at various stages. This should be done during the tuning of the model but it can already be explored outside the model starting from the stomach data. If information on prey weights is insufficient from the stomach data on Georges Bank, it is recommended to look at stomach data over a broader geographical area (e.g., NEUS). This data was previously used by Trijoulet *et al.* (2019, 2020). The other aspect on which the review team would like to comment is that, at present, the estimation of predator-prey preference is not done internally to the model. While this is a convenient simplification during the development of the estimation framework, it would be worth reconsidering in the future. Recognising all the limitations of the data, we could draw a parallel by saying that in a model like Hydra the estimation of the predator-prey size selection is as relevant as the estimation of the fleet size selectivity, and more work is recommended in that direction.

**Fit to von Bertalanffy parameters and potential bias due to stratification of sampling– MS**

The fits of the von Bertalanffy curves are done without accounting for the length stratification in aging. The parameters of the von Bertalanffy model can be sensitive to the survey design (Perreault *et al.*, 2020), with length-stratified sampling usually overestimating $L_\infty$. This will mean that the growth in Hydra will not match the growth in the population.

**Time of recruitment**

Recruitment is assumed to occur instantaneously at the beginning of every year. Depending on the species and the time step when recruitment occurs in nature and individuals appear in the data this could result in artifacts in the interactions between young individuals and predators or fleets. We recommend evaluating the impact of this assumption.

### 2.2.3 Is the input data quality and parameterization sufficient for the problem?

See above defining the problem (see section 2.2.1). Which datasets are adequate, which could be improved, and which are missing?

Show the input data as a simple chart: beginning and end of time-series, gaps, different length of time-series, spatial resolution of data.

Give information on input data pedigree/quality, references for where it comes from, whether it is survey data or comes from other model output, whether confidence intervals or other uncertainty measures are available and used in the model.

Categorize the assumptions behind modelled ecological or biological processes. Emphasize those related to species interactions (predation, competition), environmental pressures, and also fleet dynamics if needed to address the problem. If the model is spatial, how do these processes happen in space?

Is the parameterization consistent with scientific knowledge (e.g. (PREBAL) diagnostics Link (2010) for general relationships across trophic levels, sizes, etc.).

Catch data are aggregated on an annual basis regardless of the internal time step of the model under the assumption that F is constant throughout the year. An analysis of the seasonal distribution of the catches is recommended to understand if such an assumption is appropriate for the species and fisheries represented in GB (see also general section above for issues on spatial segregation of catches inside/outside the GB EPU).

Age-length data are available from surveys, and were used to estimate the transition matrix. In a context of changes in population growth parameters, it is important to ensure that the transition matrix can be updated regularly to account for changes in growth rates.

### 2.2.4 Comparison with observations

Hydra was fitted to different data sets prior to the meeting and a complete set of diagnostics was made available to the reviewers. We appreciate the work that has been put in making clear and understandable diagnostics for this model. However, given the limitations described above and the modelers agreeing on not presenting an official keyrun, the fit to observations is not commented as part of this review.

### 2.2.5 Uncertainty/sensitivity analyses

The impact of any change suggested by the review team to the current model should be documented to evaluate the trade-off between model fit and uncertainty.

As mentioned in section 2.2, sensitivity to timing of recruitment, which is currently assumed at the beginning of the year, should be evaluated.

The modelling team started directly with fitting the model to complex ecosystems. Before doing this, we suggest starting with testing the performance of the model. Here, the framework is perfect because Hydra allows for simulating data. We suggest that the team starts by simulating data with very low variance for a set of parameters and then fits this data to check that the model structure is correct and the parameters are well estimated. Thereafter, the error can be increased to test model performance further.

We would also suggest reducing the number of stocks at first as it might be difficult to understand model performance when the system is complex from the beginning, e.g., a system with one predator and two prey.

In section 2.2, we suggested estimating the vulnerability parameters within the model. To understand the properties of the model, it would be interesting to investigate the consequences of making this change by comparing the outputs and diagnostics of the model before and after change.

As mentioned in part 2.2., using one fleet per stock would solve the problem of selectivity being the same per stock as a temporary fix. Maybe this could be used as a first step before making the model more complex and integrating technical interactions.

Has uncertainty been assessed in the output of interest? Has sensitivity analysis been performed and how does it affect those outputs?

The key-run should show estimates of uncertainty in the output quantity of interest. Uncertainty analysis is best if possible to estimate confidence intervals. If not possible, list key sources of uncertainty, expected bounds on outputs based on those (possibly from sensitivity analysis)–i.e. design a sensitivity analysis to approximate the uncertainty analysis.

Specific analyses, sensitivity of key output in:

1. Retrospective analysis (5-year peel of all input data)
2. Forecast uncertainty: remove the last 3–5 years of survey indices only to see how well the model works in forecast mode, given the catch that actually happened
3. Sensitivity to stomach data and other key or low-confidence data sources
4. Sensitivity to key parameters: consumption rates, residual mortality (M1, M0)
5. Sensitivity to initial conditions

For complex models with long runtimes, simpler ways to address uncertainty may be appropriate (Kaplan and Marshall, 2016).

Best practice is to retain multiple parameterizations that meet the above criteria to allow scenario testing across a range of parameter values. Parameter uncertainty can be addressed even in complex models. A possible simple method using bounding could for example be implemented (e.g. base, low bound, and high bound productivity scenarios; Saltelli and Annoni 2010).

Could the model fit better with a different time step or bin setup? The time-step of the model is determined by the bin sizes, however the time-step of the model could be set to a reciprocal of an integer, e.g. ½ a year, as long as the time-step is not larger than the lowest time spent in any bin. The dynamics of the model change as the time-step changes, so, for the same data set, the fit of the model would be different, as could the uncertainty. It would be interesting to see how sensitive the time-step of the model is to the fit and the results of the model, and possibly if the time-step could be a parameter that is fit. This is notably relevant if the time step and bin length are made independent.

The effect of uncertainty in the growth parameter calculations could be considered.

## 2.3 Rpath

### 2.3.1 Is the model appropriate for the problem?

The Rpath model, as a mass balance model, is intended to describe the structure and flow of energy through the ecosystem, by quantifying the food web interactions for a high number of species or trophic groups. As part of the Rpath package, the Rsim family of functions are included for fitting of the historic data, estimating vulnerability parameters and simulating the dynamic of the modelled ecosystem components (Lucey *et al.* 2020). Rpath is not intended for stock assessment or provision of catch advice, but to cover the EBFM objective 1: assessing from a global perspective the consequences of certain management strategies, and particularly the impact of management strategies targeted at evaluated species on the rest of the ecosystem. The use of Rpath in the Georges Bank is planned in conjunction with other multispecies models, as well as an Operating Model in a MSE that can be used for hypothesis testing. The model is considered potentially suitable to deal with these objectives. However, the model was not finished and the fitting to the historic period could not be presented and reviewed. Hence, it was not possible to assess the performance of the complete model implementation for the Georges Bank and this review is limited to the balancing of the model.

Few general concerns were raised during the meeting regarding:

- The mismatch of the model spatial coverage with the observed distribution of a large number of populations and their commercial fisheries. Rpath has the capacity to be linked with other Rpath models developed in the Gulf of Maine and the Middle Atlantic Bight EPUs. By the time of the meeting, this extension had not been developed and could not be assessed by the reviewers, however it was considered an interesting development to address the issue of mismatch in spatial coverage.
- The datasets used when setting the Rpath parameters to obtain a balanced model for which suggestions for improvement/sensitivity assessment are presented in the section below.

### 2.3.2 Assumptions (scientific basis, computational infrastructure; adequacy of conceptual model)

Rpath is a published model and software which is an R coded version of Ecopath (Lucey *et al.* 2020). The balancing (solving of the linear equations to assess the ecotrophic efficiency) of the Rpath model is done manually according to published criteria (Link *et al.* 2010) that constrain certain parameters within bounds. Therefore, two independent implementations of balancing could lead to a different balanced model. In general, this is considered a limitation for this type of models and development of automatic optimisation procedures would be an important addition to ensure reproducibility. This should maybe be looked at in the future. Nonetheless, the balancing procedure is documented, so that each decision and its impact is tracked.

Finally, the Rpath model is checked regarding a number of qualitative emerging properties known as PREBAL diagnostics, to evaluate the realism of the system structure and flows (i.e., biomass and production per trophic levels, consistency with longevity, etc.).

Although these diagnostics are important, they are recognised as subjective and the reviewers recommend assessing the sensitivity of the fitting and projections to the initial model balancing assumptions and settings (see 2.3.5).

### 2.3.3 Is the input data quality and parameterization sufficient for the problem?

See above defining the problem. Which datasets are adequate, which could be improved, and which are missing?

Show the input data as a simple chart: beginning and end of time-series, gaps, different length of time-series, spatial resolution of data.

Give information on input data pedigree/quality, reference for where it comes from, whether it is survey data or comes from other model output, whether confidence intervals or other uncertainty measures are available and used in the model.

Categorize the assumptions behind modelled ecological or biological processes. Emphasize those related to species interactions (predation, competition), environmental pressures, and also fleet dynamics if needed to address the problem. If the model is spatial, how do these processes happen in space?

Is the parameterization consistent with scientific knowledge (e.g. (PREBAL) diagnostics Link (2010) for general relationships across trophic levels, sizes, etc.).

The Rpath model was balanced using 1981–1985 data so that the fitting could cover the period 1986-recent. During the meeting, a concern was raised regarding the fact that the balancing of the model was based on the 1981–1985 period for all parameters except the diet proportions, where the full time-series was used to allow for a larger sample size. Doing so means that the diet proportions are partly independent from the population structure in the 1980s. For instance, the early 1980s are known for having a smaller mackerel stock size than currently due to the fact that it was heavily fished at that time. Mackerel should therefore be present in lower abundance in the predator diet in the 1980s than now and the averaged proportions will show a larger proportion of mackerel in the diet. This might cause problems later in the fitting with matching the biomass or catches of mackerel. Some quick tests were recommended during the meeting to test the consequence of the diet assumption (see 2.3.5).

Parameters from EMAX for upper and lower trophic level (benthos and whales): old surveys (fifties) but not updated information

Similarly to Ecopath, a pedigree matrix is filled with regard to every parameter of the model. It has not been yet presented to the reviewers but it is a valuable tool to report on confidence and on which to base uncertainty analyses.

### 2.3.4 Comparison with observations

The model was incomplete since the fitting to the historic period (Ecosim) was not presented so it was not possible to assess how well the model fit observations.

### 2.3.5 Uncertainty/sensitivity analyses

To test the consequence of the timing inconsistency between the starting period for Ecopath and the one for the diet proportion, a sensitivity analysis was performed by changing the proportion in the predator diet to the average proportions in 1981–1985 and using these proportions directly in the balanced model. This resulted in 5 stocks that became unbalanced (EE>1).

The Ecosens procedure can be used to generate alternative parameter values for Rpath and select the ones that are balanced and allow all species to persist in projection. An analysis of these results would help understand the major factors that condition balance and persistence, and rationalize the process. It is also recommended that the robustness of the simulations (not presented at the meeting) to the Rpath parameterisation will be assessed when the model is used in projections.

## 2.4 Ensemble

The models are not currently ready to be combined in an ensemble model. The intended procedure for combining (broad sense) the models was not presented nor reviewed and only suggestions are provided in the following.

Depending on the aim of combining the models the steps towards doing so are quite different:

1. If the aim is to combine the models in a way similar to Spence *et al.* (2018), then the models should be completed, using as much information as possible. The ensemble model will describe how, with rigorously quantified uncertainty, the individual models relate to the truth. One issue we can see is the exchangeability assumption of the errors may not hold for all of the models, i.e. the surplus production model may require its own discrepancy term (Rougier *et al.* 2013).
2. If the aim is to combine the models using a model average then the individual models should be completed, using as much information as possible. Although this is often informative, it is difficult to get a formal idea of what this means. For example, it is difficult to estimate uncertainty.
3. If the aim is to do a weighted average with the weights calculated using a likelihood approach, e.g. Bayesian model averaging, then the models should be fit to the same data, which may not be possible across the three models. This is not recommended, due to the assumptions of Bayesian model averages.
4. If the aim is to compare modelling frameworks then as much of the model structures and model inputs should be shared if possible. For examples on model comparison we suggest looking at the IPCC and FishMIP and Bauer *et al.* (2019).

The differences in assumptions between Hydra and MSSPM (i.e., density dependence, species trophic interactions) should be kept in mind when they are combined, so discrepancies are interpreted in relation to these assumptions, as well as to possible differences in the input data and to structural differences. The reviewers support the intention of the modelers to use Rpath for a different purpose than the two other models (MSE loop) and not try to aggregate its outputs with those of Hydra or MSSPM.

## 2.5 References

Bauer B., Horbowy J., Rahikainen M., Kulatska N., Müller-Karulis B., Tomczak M.T., Bartolino V. (2019) Model uncertainty and simulated multispecies fisheries management advice in the Baltic Sea. PLoS ONE 14(1): e0211320. https://doi.org/10.1371/journal.pone.0211320

Bowman R.E., Michaels W.L. (1984) Part I: Examination by predator length and geographic area. In Food of seventeen species of Northwest Atlantic fish. NOAA Technical Memorandum NMFS-F/NEC-28. U.S. DEPARTMENT OF COMMERCE, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Center, Woods Hole, MA.

Gaichas S.K., Fogarty M., Fay G., Gamble R., Lucey S., Smith L. (2017) Combining stock, multispecies, and ecosystem level fishery objectives within an operational management procedure: simulations to start the conversation. ICES Journal of Marine Science 74 (2): 552-565.

Gelman A., Carlin J.B., Stern H.S., Dunson D.B., Vehtari A., Rubin, D.B. (2013). Bayesian Data Analysis (3rd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/b16018

Lucey S., Gaichas S.K., Aydin K.Y. (2020). Conducting reproducible ecosystem modelling using the open source mass balance model Rpath. Ecological Modelling 427: 109057 https://doi.org/10.1016/j.ecolmodel.2020.109057

NEFSC 2010. 49th Northeast Regional Stock Assessment Workshop (49th SAW) Assessment Report. US Dept Commerce, Northeast Fish Science Center Ref. Doc. 10-03; 383 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://nefsc.noaa.gov/publications/

Oehlert G.W. (1992). "A Note on the Delta Method". The American Statistician. 46 (1): 27-29. https://doi.org/10.1080/00031305.1992.10475842

Perreault A.M.J, Zheng N., Cadigan N.G. (2019) Estimation of growth parameters based on length-stratified age samples. Canadian Journal of Fisheries and Aquatic Sciences. 77 (3): 439-450. https://doi.org/10.1139/cjfas-2019-0129

Rougier J., Goldstein M., House L. (2013) Second-Order Exchangeability Analysis for Multimodel Ensembles, Journal of the American Statistical Association, 108:503, 852-863, https://doi.org/10.1080/01621459.2013.802963

Spence, MA, Blanchard, JL, Rossberg, AG, *et al.* A general framework for combining ecosystem models. Fish Fish. 2018; 19: 1031– 1042. https://doi.org/10.1111/faf.12310

Spence M.A., Bannister H.J., Ball J.E., Dolder P.J., Griffiths C.A., *et al.* (2020) LeMaRns: A Length-based Multi-species analysis by numerical simulation in R. PLOS ONE 15(2): e0227767 https://doi.org/10.1371/journal.pone.0227767

Trijoulet V., Fay G., Curti K., Smith B., Miller T.J. (2019). Performance of multispecies assessment models: insights on the influence of diet data. ICES Journal of Marine Science 76(6): 1464-1476 https://doi.org/10.1093/icesjms/fsz053

Trijoulet V., Fay G., Miller T.J. (2020) Performance of a state-space multispecies model: what are the consequences of ignoring predation and stochastic processes in stock assessments? Journal of Applied Ecology 57: 121-135 https://doi.org/10.1111/1365-2664.13515

# Annex 1: List of participants and Resolution

| Name | Institute | Country (of institute) | Email |
|---|---|---|---|
| Valerio Bartolino (chair) | Swedish University of Agricultural Sciences, Department of Aquatic Resources, Institute of Marine Research | Sweden | valerio.bartolino@slu.se |
| Michael Spence (chair) | Centre for Environment, Fisheries and Aquaculture Science (CEFAS) | United Kingdom | michael.spence@cefas.gov.uk |
| Sarah Gaichas | NOAA Northeast Fisheries Science Center, Woods Hole | USA | Sarah.Gaichas@noaa.gov |
| Sean Lucey | NOAA Northeast Fisheries Science Center, Woods Hole | USA | Sean.Lucey@noaa.gov |
| Vanessa Trijoulet | DTU-Aqua | Denmark | vtri@aqua.dtu.dk |
| Morten Vinther | DTU-Aqua | Denmark | mv@aqua.dtu.dk |
| Alfonso Perez | Instituto Español de Oceanografia IEO | Spain | alfonso.perez@ieo.csic.es |
| Chris Legault | NOAA Northeast Fisheries Science Center, Woods Hole | USA | chris.legault@noaa.gov |
| Sigrid Lehuta | IFREMER | France | sigrid.lehuta@ifremer.fr |
| Kiersten Curti | NOAA Northeast Fisheries Science Center, Woods Hole | USA | kiersten.curti@noaa.gov |
| Gavin Fay | School for Marine Science and Technology, University of Massachusetts Dartmouth | USA | gfay@umassd.edu |
| Andy Beet | NOAA Northeast Fisheries Science Center, Woods Hole | USA | andrew.beet@noaa.gov |
| Howard Townsend | NOAA Fisheries, Silver Spring | USA | Howard.Townsend@noaa.gov |
| Robert Gamble | NOAA Northeast Fisheries Science Center, Woods Hole | USA | Robert.Gamble@noaa.gov |
| Ron Klasky | NOAA Fisheries, Silver Spring | USA | ronald.klasky@noaa.gov |
| Christina Perez | School for Marine Science and Technology, University of Massachusetts Dartmouth | USA | mperez12@umassd.edu |
| Maxwell Grezlik | School for Marine Science and Technology, University of Massachusetts Dartmouth | USA | mgrezlik@umassd.edu |
| Floor Soudijn | Wageningen Marine Research | Netherlands | floor.soudijn@wur.nl |
| Alan Baudron | Marine Scotland Science | United Kingdom | Alan.Baudron@gov.scot |
| Andrea Belgrano | Swedish University of Agricultural Sciences, Department of Aquatic Resources, Institute of Marine Research and Swedish Institute for the Marine Environment (SIME) | Sweden | andrea.belgrano@slu.se |
| Jamie Tam | Bedford Institute of Oceanography · Oceans and Ecosystem Sciences Division | Canada | Jamie.Tam@dfo-mpo.gc.ca |
| Jacob Bentley | Natural England, London | United Kingdom | Jacob.Bentley@sams.ac.uk |
| Christopher Griffiths | Swedish University of Agricultural Sciences, Department of Aquatic Resources, Institute of Marine Research | Sweden | christopher.griffiths@slu.se |

| | | | |
|---|---|---|---|
| Francisco De Castro | Fisheries & Aquatic Ecosystems. AFBI | United Kingdom | francisco.decastro@afbini.gov.uk |
| Carolyn McKeon | Agri-Food and Biosciences Institute (AFBI) | United Kingdom | mckeonc2@tcd.ie |
| Miriam Puts | Thünen Institute of Sea Fisheries | Germany | miriam.puest@thuenen.de |
| Gustav Delius | Department of Mathematics, University of York | United Kingdom | gustav.delius@york.ac.uk |
| Michael Thomson | Centre for Environment, Fisheries and Aquaculture Science (CEFAS) | United Kingdom | michael.thomson@cefas.gov.uk |
| James Martindale | Centre for Environment, Fisheries and Aquaculture Science (CEFAS) | United Kingdom | james.martindale@cefas.co.uk |
| Alex Holdgate | Centre for Environment, Fisheries and Aquaculture Science (CEFAS) | United Kingdom | alex.holgate@cefas.gov.uk |
| Mikaela Potier | IFREMER, Agrocampus Ouest | France | mikaela.potier@agrocampus-ouest.fr |
| Liesa Celie | The Flanders Research Institute for Agriculture, Fisheries and Food | Belgium | liesa.celie@ilvo.vlaanderen.be |
| Robert Thorpe | Centre for Environment, Fisheries and Aquaculture Science (CEFAS) | United Kingdom | robert.thorpe@cefas.gov.uk |
| Marie Savina Rolland | IFREMER | France | Marie.Savina.Rolland@ifremer.fr |
| Marianne Robert | IFREMER | France | Marianne.Robert@ifremer.fr |
| Roweena Patel | University of Reading | United Kingdom | r.patel@pgr.reading.ac.uk |
| Thomas Del Santo O’Neill | Queen Mary, University of London | United Kingdom | delsantooneillthomas@gmail.com |
| Maria Ching Villanueva | IFREMER | France | Ching.villanueva@ifremer.fr |
| Jamie Lentin | Shuttle Thread Limited | United Kingdom | lentinj@shuttlethread.com |

# WGSAM Resolution

The **Working Group on Multispecies Assessment Methods** (WGSAM), chaired by Valerio Bartolino Sweden; and Michael Spence, UK, will work on ToRs and generate deliverables as listed in the Table below.

| | MEETING DATES | VENUE | REPORTING DETAILS | COMMENTS (CHANGE IN CHAIR, ETC.) |
|---|---|---|---|---|
| Year 2022 | 10-14 October | Woods Hole, USA | Reports on keyrun reviews to be provided after each review is complete | Incoming co-chair: Michael Spence (UK)<br>Outgoing co-chair: Sarah Gaichas (USA) |
| Year 2023 | 9-13 October | tbd | Reports on any keyrun reviews that are completed | |
| Year 2024 | | | Final report by Date to SCICOM | |

**ToR descriptors**

| TOR | DESCRIPTION | BACKGROUND | SCIENCE PLAN CODES | DURATION | EXPECTED DELIVERABLES |
|---|---|---|---|---|---|
| a | Regional updates: Review further progress and deliver key updates on multispecies modelling and ecosystem data analysis contributing to modeling throughout the ICES region | This ToR acts to increase the speed of communication of new results across the ICES area | 5.1; 5.2; 6.1 | 3 years | Report on further progress and key updates.<br>Review and collaborate with appropriate EGs to revise sections on "species interactions" in the Fisheries Overviews |
| b | Key-runs: Parametrisation of multispecies and ecosystem key-run models for different ICES regions. This includes standard update (limited to inclusion of recent data), extensive update (incl. new data and processes), and new key-runs. | Key-runs are models checked against high quality criteria, which are developed to contribute to a variety of operational objectives as part of the ICES advice, i.e. provide information on natural mortality for inclusion in single species assessments, estimates of multispecies reference points, large operating ecosystem models for MSE, etc. | 5.1; 5.2; 6.1 | 3 years | Report on output of multispecies models including stock biomass and numbers and natural mortalities for use by single species assessment groups and external users. |
| c | Skill assessment: Establish and apply methods to assess the skill of multispecies | This work is aimed at assessing the performance of models intended for | 5.1; 6.1; 6.3 | 3 years | Report on technical requirements for cross-models standardisation and comparison. |

| | | | | | |
|---|---|---|---|---|---|
| | models intended for operational advice | strategic or tactical management advice. Evaluation will require work towards standardisation for cross-model comparison. This ToR will also deal with evaluation of methods for model calibration and data weighting in the context of multispecies modelling. | | | Manuscript(s) on skill assessment of wide array of multispecies models based on a large simulation study. |
| d | Multi-model advice: Evaluate methods for generating advice by comparing and/or combining multiple models | This work is aimed at addressing structural uncertainty in advice arising from multiple models | 5.1; 6.1; 6.3 | 3 years | Report on methods for comparing models and for constructing model ensembles.<br>Report on case examples from both simulation testing and real studies |
| e | MSE: Evaluate methods and applications for multispecies and ecosystem advice, including evaluation of management procedures and estimation of biological reference points under the uncertainties of climate change. | This ToR looks for multispecies and ecosystem approaches to understand the resistance and resilience of ecosystems to a warming environment and to perturbations related to the effects of climate change. Through the use of simulations, alternative management strategies and exploitation regimes can be evaluated for robustness to uncertainties related to climate change. | 2.5; 5.2; 6.1 | 3 years | Review methods to evaluate populations and ecosystem resilience.<br>Review of methods for management strategy evaluation which incorporate the effects and uncertainties of climate change |

## Summary of the Work Plan

| | |
|---|---|
| Year 1 | All ToRs, update keyrun Baltic Sea (coupled with data preparation workshop for the Baltic Sea benchmark), keyrun Georges Bank multi-model (dedicated workshop) |
| Year 2 | All ToRs |
| Year 3 | All ToRs |

**Supporting information**

| | |
|---|---|
| Priority | The current activities of this Group will lead ICES into issues related to the ecosystem<br>effects of fisheries under multiple sources of uncertainties incl. climate change. The activities will provide information (e.g., natural mortality estimates, performance of<br>indicators, multispecies reference points) and tools (e.g., multi-model ensembles, keyrun models) valuable for the implementation of an integrated advice and the application of a precautionary approach in several North Atlantic ecosystems. Consequently, these activities are considered to have a high priority. |
| Resource requirements | The research programmes which provide the main input to this group are already underway, and resources are already committed. The additional resource required to undertake additional activities in the framework of this group is negligible. |
| Participants | The Group is normally attended by some 20–25 members and guests. Expertise in ecosystem dynamics, trophic interactions, modelling and fish stock assessment from across the whole ICES region. |
| Secretariat facilities | Standard EG support. |
| Financial | No financial implications. |
| Linkages to ACOM and groups under ACOM | ACOM, assessment Expert Groups. |
| Linkages to other committees or groups | WGMIXFISH, WGDIM, WGBIFS, IBTSWG, WGECO, all IEASG groups, WKCLIMAD. |
| Linkages to other organizations | None |

# Annex 2: Stock Annex for the ICES Eastern Baltic Sea SMS configuration

| | |
|---|---|
| Working Group | Working Group on Multispecies Assessment Methods (WGSAM) |
| Date | October 2022 (WGSAM 2022 meeting) |
| Predatory species | Cod |
| Prey species | Assessed species: Herring, Sprat |
| Stock Assessor | Morten Vinther |

## Summary

The keyrun uses the SMS model (Lewy and Vinther, 2004) which is a stock assessment model including biological interactions estimated from a parameterised size-dependent food selection function. The model is formulated and fitted to observations of total catches, survey cpue and stomach contents for the Eastern Baltic Sea (ICES Sub-divisions 25-32, excluding the Gulf of Riga). Parameters are estimated by maximum likelihood and the variance/covariance matrix is obtained from the Hessian matrix.

In the present SMS analysis, cod is a predator, and herring and sprat are preys. The population dynamics of cod were estimated outside the model by ICES WGBFAS, whereas keyruns before 2019 estimated cod stock size and cod cannibalism within the SMS.

Substantial changes of input data were part of the 2019 keyrun, but the 2022 keyrun is mainly an addition of stock assessment data of the last three years and a small correction of the food ration calculation.  The 2022 estimated predation mortalities (M2) are consistent with the M2 values from the previous keyrun in 2019.

### 2019 keyrun

A keyrun for the Eastern Baltic Sea SMS model, including data for the period 1974–2018 was produced at the 2019 WGSAM. This keyrun replaced the 2012 keyrun. The 2019 keyrun included revisions and updates of the input data. Major modifications were that cod is treated as an external predator and the use of newly available data on cod stomach contents sampled mainly by the Latvian Institute.

SMS was updated with the most recent data from WGBFAS 2019, i.e. data for Herring in subdivisions 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea) and for Sprat in subdivisions 22–32.

Due to age reading problems for cod in the eastern Baltic, ICES has since 2019 applied an age-length based analytical assessment with the Stock Synthesis model (SS3). Natural mortality of cod is estimated within the SS3 model. Without input data by ages, and with estimated high and time variable natural mortality, SMS is no longer able to estimate cod stock numbers and predation mortality estimates on cod due to cannibalism. Instead, cod is now considered as an "other predator" where stock number and size distribution are assumed to be known without errors. Population numbers and size distributions were extracted from the SS3 output.

Consumption (food ration) of cod was revised to reflect the most recent knowledge of evacuation rates and temporal trends in cod consumption rates.

Diet data for cod were substantially extended by including the stomach content data from the EU Stomach Tender. This addition of data did not change predation rates on herring and sprat substantially, but increased the weighting of the stomach data in the model likelihood, indicating a higher quality of stomach data compared to the previously used data.

### 2022 keyrun

The 2022 keyrun includes revisions and updates of the input data (1974-2021). New assessment input data (e.g. catch at age numbers, mean weight and survey indices) , 2019-2021, from WGBFAS were included. Cod diet data were recalculated from the same stomach data observations (1974-2014) as used in the 2019 keyrun. Cod food rations were also recalculated, to correct for seasonal variation. The model configuration was changed slightly for Fishing mortality. Overall these changes are not considered substantial, and the estimated keyrun results are close to the results from the 2019 keyrun.

# Model description

The SMS model (Lewy and Vinther, 2004) is a stock assessment model including biological interactions estimated from a parameterised size-dependent food selection function. The model is formulated and fitted to observations of total catches, survey cpue and stomach contents for the main stocks in the Baltic Sea. Parameters are estimated by maximum likelihood and the variance/covariance matrix is obtained from the Hessian matrix.

The following predator and prey stocks are available:

- External predator: cod;
- Prey: herring and sprat

The population dynamics of herring and sprat are estimated within the model.

A detailed description of the model can be found in Appendix 1.

# Input data

The description of input data is divided into four main sections:

**Analytical assessment stocks**: Stocks for which analytical age-based assessments are done by ICES or can be done from data available from ICES. Data input are similar to those applied by ICES “single-species” assessments used for TAC advice, with some additional data.

**External predator stocks**: Stocks for which stock numbers are assumed known and given as input to SMS.

**Diet and ration data**: Diet data and food ration data for all predators (analytical stocks and external predators) derived from observed stomach contents data.

**Additional data**: Miscellaneous data.

## Analytical assessment stocks

This group of stocks includes:

1 ) Herring;
2 ) Sprat;

“Single-species” input data, by default given by quarterly time steps, include

- Catch-at-age in numbers (SMS input file canum.in);
- Proportion of the catch-at-age landed, assumed 100% (file proportion_landed.in);
- Mean weight-at-age in the catch (file weca.in);
- Mean weight-at-age in the stock (file west.in);
- Proportion mature-at-age (file propmat.in);
- Proportion of M and F before spawning (file proportion_M_and_F_before_spawning.in);
- M, single-species natural mortality-at-age (file natmor.in);
- Survey catch-at-age and effort (file fleet_catch.in).

SMS uses quarterly time steps, so input catch data should preferably also be given by quarter which is also the case for the Baltic Sea.

**Table 2.1.1. Overview of "dynamic" fish stocks used in SMS and their basis from ICES single-species advice.**

| Species | SMS | | ICES Assessment | | | |
|---|---|---|---|---|---|---|
| | Species code | Max age | Stock area | First year | Age range (data) | time step |
| Herring | HER | 8+ | SD 25–29 and 32, excluding the Gulf of Riga (central Baltic Sea) | 1974 | 1–8+ | Quarter |
| Sprat | SPR | 7+ | SD 22-32 | 1974 | 1–8+ | Quarter |

Discarding is considered to be negligible for both stocks.

### Herring

#### Catch numbers at age

Data for 2019-2021 were copied from Excel sheets stored at the ICES WGBFAS SharePoint. ICES WGBFAS provided quarterly catch-at-age number and mean weights for herring for the period 2002-2018. The full data series are not presented in the WGBFAS report, but were kindly made available by ICES stock assessor Tomas Gröhsler.  Catch at age for the period before 2002 data were copied from the 2012 SMS keyrun.

#### Mean weight at age

WGBFAS assumes that mean weight at age in the sea is the same as mean weight at age in the catch. This assumption is fairly unbiased for older fish even though fisheries may be concentrated in areas (southern part of the EB) with the largest individuals. Mean weight at age in the catch for the youngest fish is higher than the mean weight in the sea as these size classes are not fully selected in the fishery.  The mean weight at age as used by WGBFAS (Figure 2.1-1) shows a clear temporal trend with a decreasing mean weight in the period 1974-2000.

The quarterly mean weight at age data from WGBFAS (2002-2021) combined with the 2012 keyrun data for the period 1974-2001 are presented in Figure 2.1-2 for the youngest ages 0 and 1. It is clearly seen that the mean weights for age 1 in quarter 2 do poorly link to quarter 1 and not at all to quarter 3.

It is assumed that the mean weight in the sea are the same as the observed mean weight in the catch. However, when calculation the mean weight at age in the sea, the observed mean weight at age in the catch for age 1 in quarter 2 was discarded and substituted by the average of the observed mean weight at age in quarter 1 and 3 (Figure 2.1-3). As the observed mean weights for the ages 0-1 are highly variable, the smoothed values, ages 0 and 1, were finally used as mean weight in the sea.

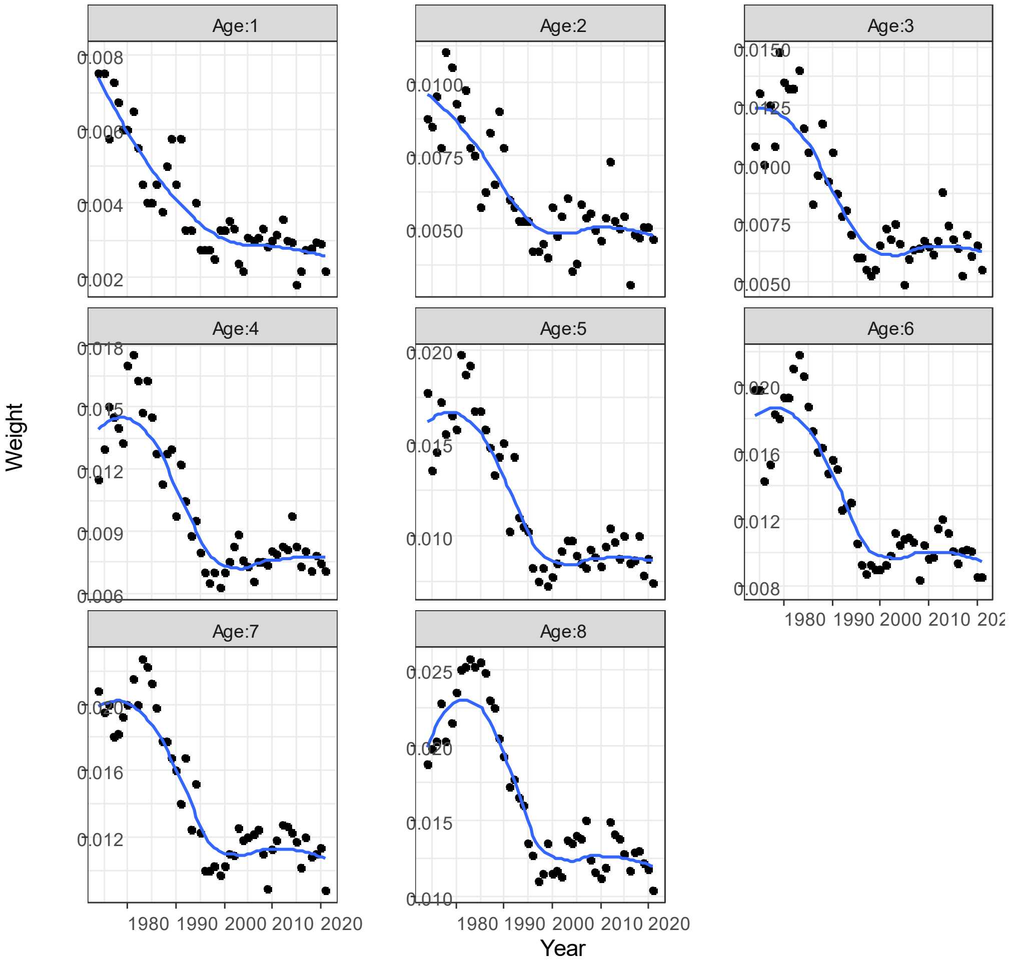


Figure 2.1-1. Herring mean weight at age in the catch (and in the sea) as used by WGBFAS. Dots show data points and the blue line is a loess smoother.

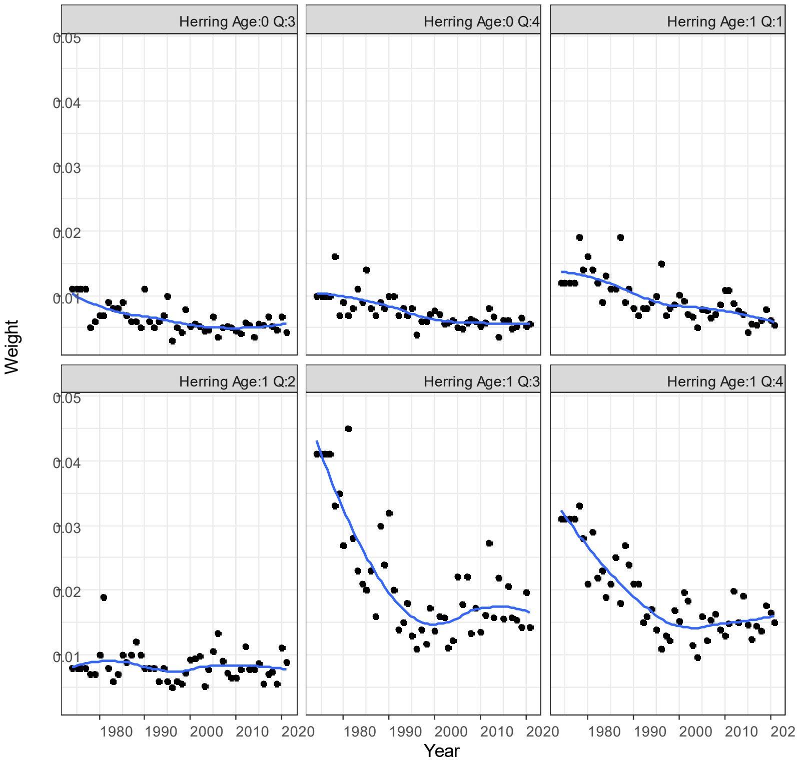


Figure 2.1-2. Quarterly herring mean weight at ages 0 and 1 in the catch as provided by WGBFAS. Dots show data points and the blue line is a loess smoother.

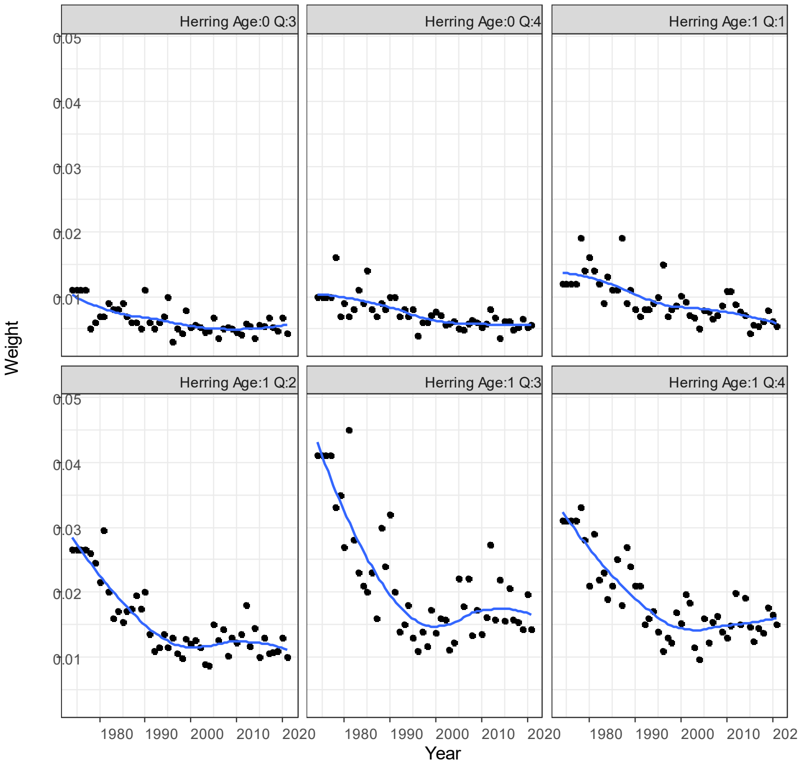


Figure 2.1-3. Quarterly herring mean weight at ages 0 and 1 in the sea as used by SMS. Dots show data points and the blue line is a loess smoother.

#### Other biological data

Proportion mature and M (used for "single species" SMS) at age data are copied from single-species data. WGSAM 2019 decided to use M1 at 0.025 per quarter for all ages. The 2012 key-run applied 0.05, but for consistency with herring in the North Sea this was changed to 0.025 in 2019 and this is also the value used in the 2022 keyrun.

#### Survey data

Survey data are copied from the previous keyrun and the ICES single-species assessment.

| SMS name | Years | Ages | alfa and beta | Source |
|---|---|---|---|---|
| Herring_Acoustic_May | 1982-1996 | 1–8 | 0.2-0.7 (Q2) | 2012 keyrun |
| Herring_Acoustic_BIAS | 1991-2021 | 1–7 | 0.0-0.3 (Q3) | WGBFAS, 2022 |

### Sprat

#### Catch at age

Data for 2019-2021 were copied from Excel sheets stored at the ICES WGBFAS SharePoint. Quarterly catch-at-age number and mean weights for sprat, 1998-2018, were provided by ICES WGBFAS. The full data series are not presented in the WGBFAS report, but were kindly provided by Tomas Gröhsler. Older quarterly catch at age data were copied from the 2012 SMS keyrun.

#### Mean weight at age

WGBFAS assumes that mean weight at age in the sea is the same as mean weight at age in the catch. This assumption is probably unbiased for older fish even though fisheries may be concentrated in areas (south-western part of the EB) with the largest individuals. Mean weight at age in the catch for the youngest fish is probably higher than the mean weight in the sea as these size classes are not fully selected in the fishery. The mean weights at age as used by WGBFAS (Figure 2.1-4) show a clear temporal trend with a peak in mean weight around 1987 followed by a decrease until around 2003.

The quarterly mean weight at age in the catch from WGBFAS (1998-2018) combined with the 2012 keyrun data for the period 1974-1978 are presented in Figure 2.1-5 for the youngest ages 0 and 1. It is clearly seen that the mean weights for age 0 and age 1 in quarter 1 and 2 are highly variable from one year to the next. The same can be said about age 1 in quarter 3 and 4, but these quarters follow better the overall trend presented for the WGBFAS data (Figure 2.1-4). Due to the high (observation) variation in catch mean weights for ages 0-1, the smoothed values were used as mean weight at age in the sea (Figure 2.1-5) by SMS.

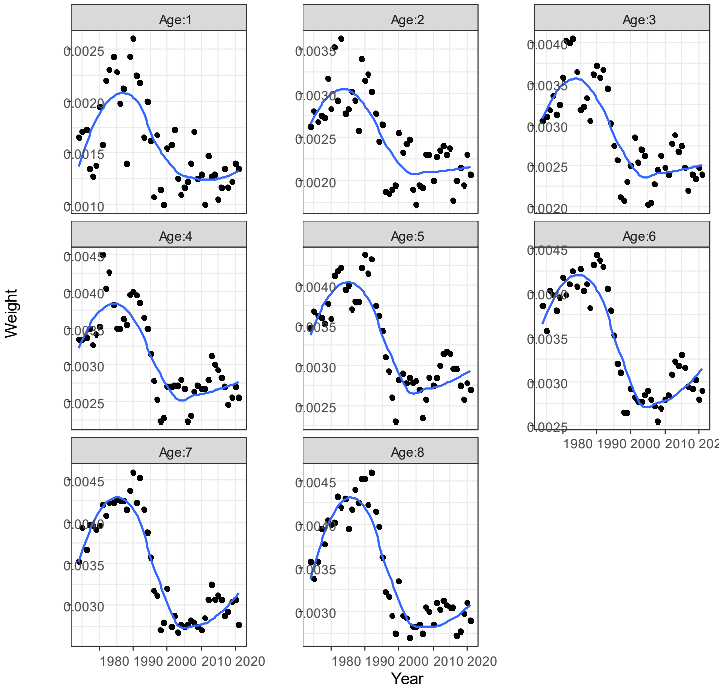


Figure 2.1-4. Sprat mean weight at age in the catch (and in the sea) as used by WGBFAS. Dots show data points and the blue line is a loess smoother.

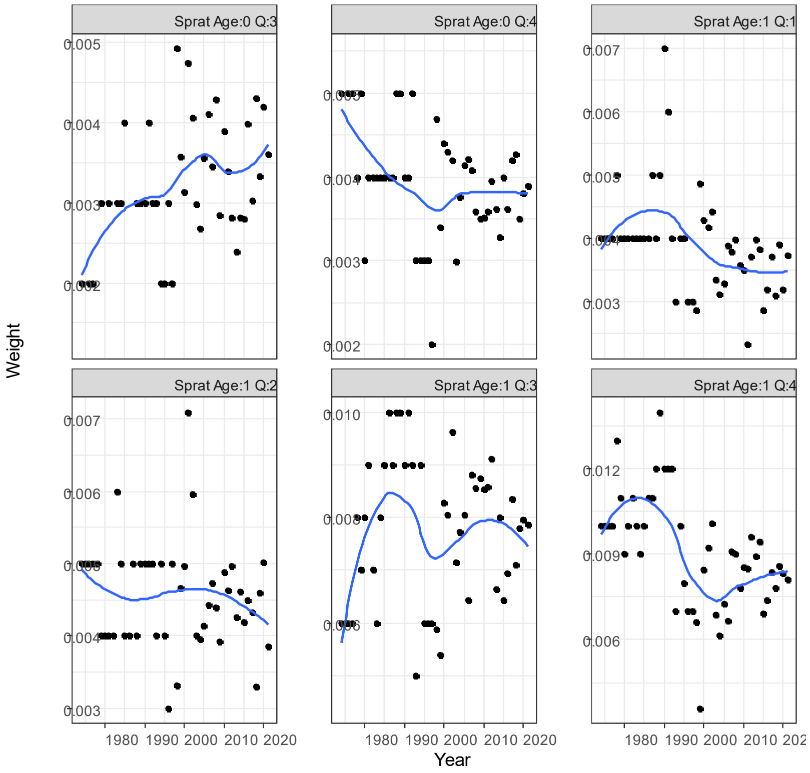


Figure 2.1-5. Quarterly sprat mean weight at ages 0 and 1 in the catch as provided by WGBFAS. Dots show data points and the blue line is a loess smoother.

#### Survey data

Survey data are copied from the single-species assessment (survey 1–3).

| | NAME | YEARS | AGES | ALFA AND BETA | SOURCE |
|---|---|---|---|---|---|
| 1 | Int acoustic in Oct. | 1991-2018 | 1–7 | 0.0–0.1 (Q3) | WGBFAS 2022 |
| 2 | Int_acoustic_in_May | 2001–2018 | 1–7 | 0.25–0.50 (Q2) | WGBFAS 2022 |
| 3 | LAT_RUS_acoustic | 1992-2018 | 1–1 | 0.0-0.0 (Q1) | WGBFAS 2022 |

#### Biological data

Proportion mature and M at age data are copied from single-species data. M1 is assumed to be 0.05 per quarter for all ages.

### External predators

In the 2019 keyrun, cod was for the first time in the Baltic SMS treated as an "external predator". This means that the stock numbers are given by input, extracted from the ICES WGBFAS, Stock-Synthesis 3 (SS3) assessment for the stock. The SS3 assessment provides cod stock numbers and mean weight by 2-cm length classes for the main length classes. These data were aggregated into length classes used by SMS stomach contents data.

The SS3 assessment output is quite different from the previous age-based assessment from the 2012 key-run (Figure 2.2-1). The SS3 assessment estimates much higher stock numbers for ages 1-3 compared to the SMS estimate, and higher stock numbers for oldest cod when the stock size peaked. The SS3 results from the 2022 assessment are quite similar to the 2019 results.

SS3 and SMS use different mean weight at age, so the difference in biomass, quarter 1, (Figure 2.2-2) becomes smaller than when stock numbers were compared. Total biomass and biomass of the larger cod are estimated higher in the SS3 assessments (Figure 2.2-3) such that the amounts of food eaten and predation mortality become higher when the cod estimate from SS3 is used. The SS3 results from 2022 are similar to the 2019 results.

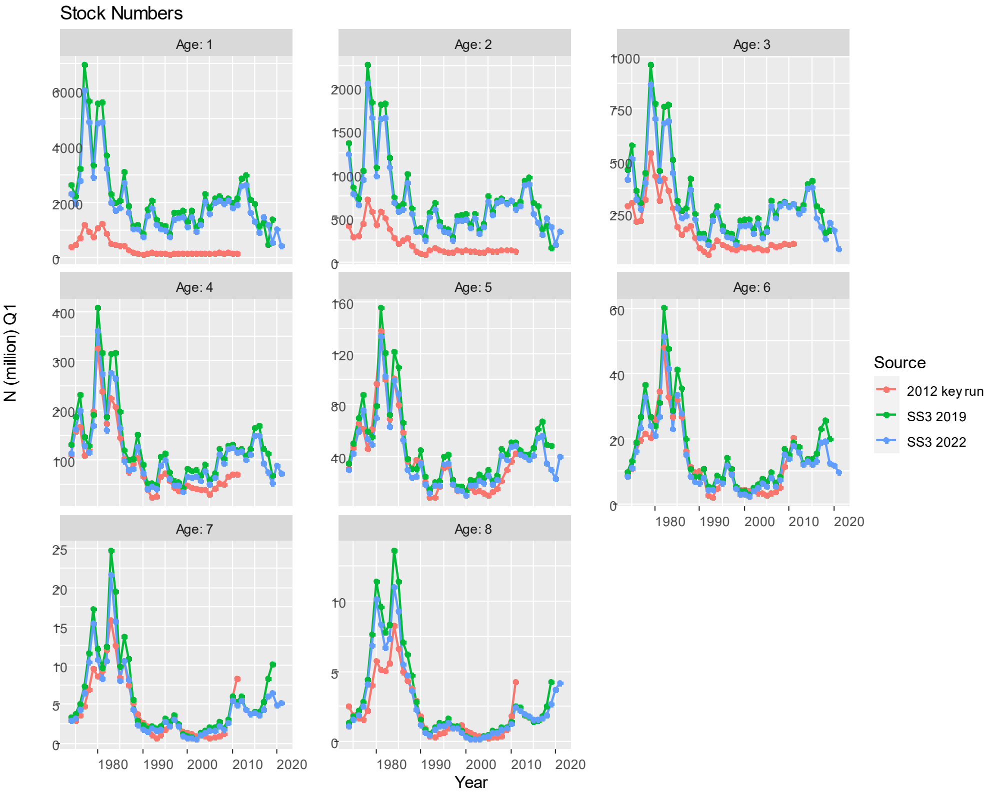


**Figure 2.2-1. Stock numbers at age of cod estimated by the 2012 key-run and by the ICES SS3 assessments in 2019 and 2022.**

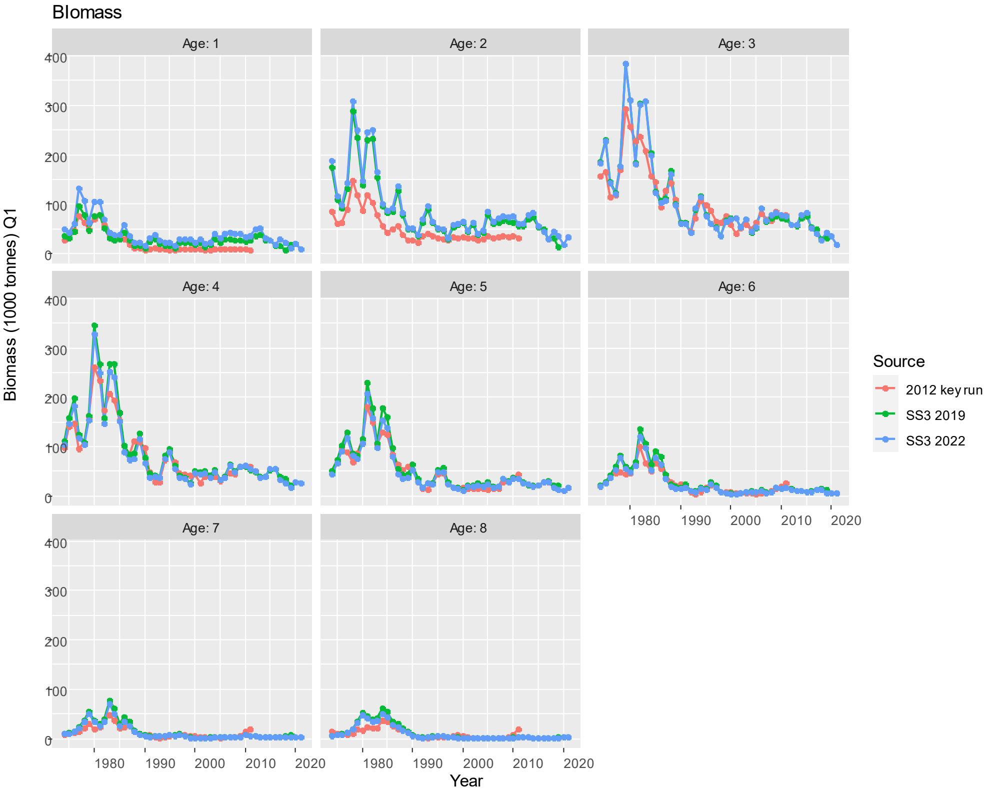


**Figure 2.2-2. Biomass at age of cod estimated by the 2012 key-run and by the ICES SS3 assessments in 2019 and 2022.**

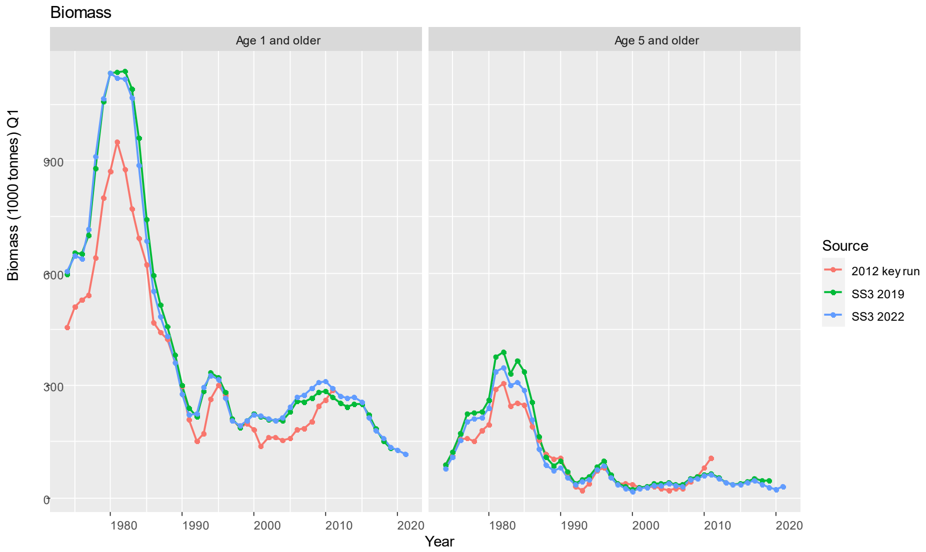


**Figure 2.2-3. Cod biomass for age 1 and older (left panel) and of age 5 and older (right panel) estimated by the 2012 key-run and by the ICES SS3 assessments in 2019 and 2022.**

## Diet and ration data

### Fish stomach data

Two major datasets of cod stomach contents are available:

- "Old": International sampled stomach content data, 1977-1992, Individual stomachs were pooled by cod size class before analysis. The recorded sizes of both predator and are given by wide size classes, e.g. sprat by the size classes 5-10-15 cm, for the oldest data in the time series.
- "New": Individually compiled stomach sampled by mainly Latvia in the period (1963) 1974-2014. Predator and prey sizes are by cm or mm.

#### "old" pooled stomach data

An international database of Baltic cod stomach contents contains data from 62 427 cod collected during 1977–1994. The collation of national stomach content data sets into one set for multispecies assessment has mainly been done by DIFRES (now DTU Aqua) and the results were published in ICES papers (e.g. ICES 1991/J:30; ICES 1989/J:2; ICES 1990/Assess:25 and ICES 1993/J:11). The stomach contents data are available at "exchange format" from ICES (www.ices.dk).

The "old" data stomach contents data are recorded by year, quarter, predator, predator size class, prey and prey size class. Most stomachs were pooled within a haul and predator size class before analysis, such that diet data from individual fish are scarce. For part of the time series, data were only provided (pooled) by country and ICES subdivision (SD), such that the variation between hauls could not be analysed.

#### "new" individually sampled stomach data

More than one hundred thousand stomachs of cod from the Eastern Baltic Sea have been sampled by trawling between 1963 and 2014, by mainly the Latvian institute. Sampling (Figure 2.3-1) covered the distributional area of the Eastern Baltic cod population (Bagge, 1994) except in the period 1995 to 2004, where sampling was limited to the north-eastern part. Stomach contents are provided by individual fish. Prey items in the stomachs were recorded at the highest possible taxonomic resolution with total mass, and, where identifiable, number of individuals and lengths per prey taxon. Prey sizes are given by mm or cm. Predator length was also recorded and in later years also predator weight (Huwer et al., 2014;). The stomach data are available at ICES (www.ices.dk), however the ICES download facility has not been working for the last 2 years!.

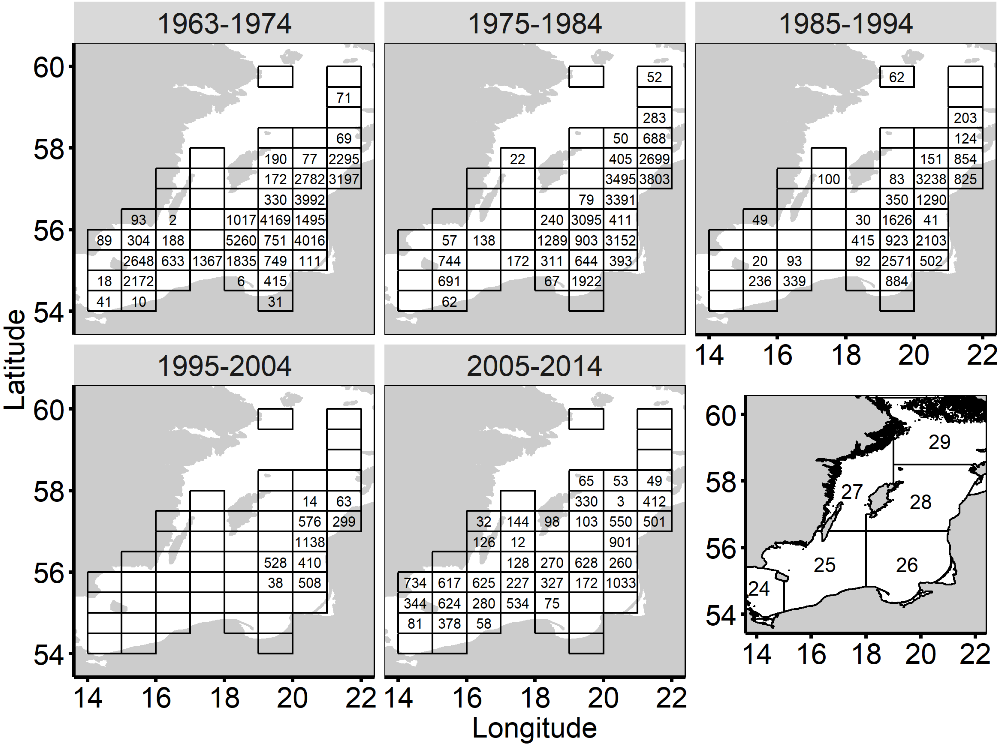


**Figure 2.3-1. ICES sub-divisions (bottom right panel) and stomach sampling coverage: number of *Gadus morhua* stomachs by ICES statistical rectangle for each period specified on top of each panel. Source Neuenfeldt et al., 2020.**

#### Pre-processing of data

In the Latvian data before 2000, prey weight is not given per prey item, but per prey species and digestion stage. To assign weight to each prey item, a length weight relation was first made for each prey species and digestion stage, based on stomach data from predators with only one prey item of a given species. Secondly, these length weight relations were used to assign a weight to each prey with size information. The sum of these weights cannot exceed the total recorded prey weight for the individual stomach. If the sum exceeded the total prey weight (of both sizes and un-sized preys), the mean weight of the sized preys were downscaled and prey items with no size information was removed. If the sum was smaller than the recorded total weight, and the stomach included preys without size information, the difference in weight was assigned to the prey with no size information.

#### Compilation of stomach contents data into diet

The compilation of the individual stomach samples from trawl hauls into average diet of the Eastern Baltic Sea basically follows the technique given by ICES (1993). The compilation of stomachs contents for the 2022 keyrun was done using the FishStomach R-package (available from https://github.com/MortenVinther/FishStomachs).

The compilation of stomach contents data into diet data includes the following steps:

**1. Read and check data from agreed exchange format**

This step reads data from the ICES exchange format for stomach contents. The "new" data were initially transformed into a format similar to the ICES format as part of the paper Neuenfeldt et al. 2020. The function "read_exchange_data" from the FishStomach package was applied for reading the data in ICES exchange format; more information about the function can be found in the man-pages and vignettes of FishStomachs)

**2. Pre-processing of "new" data**

Weight by prey items were allocated where total prey weight was the only weight available – see previous section. FishStomach function prey_w_from_pooled_weight was applied.

**3. Non assessment species into "other" prey species**

Prey species other than herring, sprat and the unidentified mix of those ("Clupeidae" and "Clupeoides") were allocated to "other" species. As cod is an "other predator" cod preys are also included in "other". Function group_prey_species was used for this step.

**3. Assign size classes for predators and preys**

The length of predators of preys were from the "new" data were recorded in mm or cm the sizes were assigned to sizes classes using the length classes provided in Table 2.4-2. A wider size range, based on the initial applied size classes (e.g. 5-10-15 cm for preys) were applied for the "old" data set, or for combinations of the "old" and "new" data. Functions put_size_class_on_prey and put_size_class_on_predators were used.

**4. Bias correct to take into account regurgitated stomachs within a sample unit**

The "old" data compiled as a pooled stomachs, e.g. stomach contents from all 40-50 cm cod in a haul. For each stomach pool, data include the information on the number of a) empty stomachs; b) stomach with skeleton remains only; c) stomach with food and d) stomach with food, but regurgitated. In most cases, stomachs within a haul are pooled at the time of sampling for each predator size class. Only stomach contents from the feeding, non-regurgitated stomachs were recorded and later bulked to save time. In the calculation of the average stomach content, it was assumed that the regurgitated stomachs had similar stomach content as the (valid) feeding fish. The "new" data set does not include information on the presence of regurgitated stomachs, and no corrections were made. Function bias_correct_regurgitated was used.

**5. Aggregate stomachs contents within sample_id and size classes**.

Average stomach contents by prey species and size class are calculated for all stomachs within the same sample_id (by haul) and predator size class. Function aggregate_within_sample was used.

**6. Stratification for allocation of missing data, and calculation of population diet**

Strata are used to group data for calculation of mean proportions of e.g. the size distribution of a given prey, and for raising estimated stomach contents to population diet. Three strata were defined as: 1. stratum_time: year and quarter; 2. stratum_sub_area; ICES subdivision (i.e, SD25, SD26 and SD 28) and 3. stratum_area: ICES subdivision, same as above. Function add_strata was used.

**7. Allocation of partly identified prey species to known prey species distribution within strata.**

Non species identified Clupeoides were allocated to the proportion of identified herring and sprat, using a series of allocation keys calculated by groups of defined from combinations of the defined strata. First allocation keys are derived from data derived by the most "local" area and time period (e.g. within the same predator size class, year-quarter combination and SD), but if sufficient data could not be found for construction of allocation keys, the search area for data was expanded e.g. using data from the same year irrespective of quarter.

**8. Allocation of identified prey species with missing size information to known prey sizes within strata.**

Size information was allocated to prey items without a size measurement using the same approach as in step 7, with use of "local" allocation data as far as possible. Function redist_unidentified_prey_lengths was used.

**9. Calculation of population diet.**

The population diet should reflect the average diet of the predators within the model area. First, the average stomach content in weight of the individual prey and prey size classes was calculated by ICES SD as a weighted mean of the available samples within the SD, weighted by the number of fish in the sample. The absolute weights were then transformed into relative weights (proportions) of the individual prey and prey size classes. Finally the population diet was calculated as a weighted average of the prey proportion by the three SDs, weighted by the number of stomachs sampled within the SD. Function calc_population_stom was used.

This approach assumes that sampling of the stomachs largely follows the spatial abundance of the predator. This has however, not been the case in all years (see Figure 2.3-1).

Diet data for the 2019 keyrun were compiled following the same steps, using an early version of the same R-code as included in FishStomach, however, the diet dataset from the 2019 and the 2022 keyrun came out different (Figure 2.3-2). Differences are due to both a change in methodology (and coding) and in the approach taken to allocate partly identified prey items to fully identified preys or to the group of other prey. The raising procedure of stomach contents to population diet data has also changed.

Due to time limitations, the diet of the population was estimated based on the assumption that the observed stomach contents give an unbiased estimate of the diet. This is in contrast to the assumptions used for the estimation of food ration, as outlined in section below, where it is assumed that the stomach evacuation rate (and thereby the food intake) depends on e.g. the energy contents of the stomach contents (as described by FishStomach function bias_correct_energy_etc ).

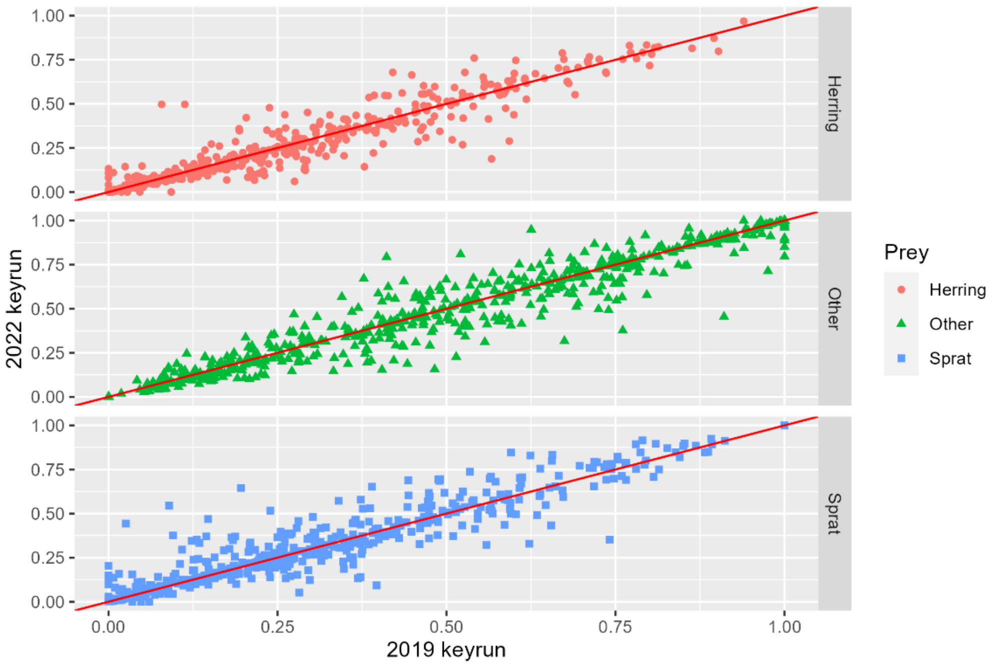


**Figure 2.3-2 Comparison of prey proportions in the diet data used in the 2019 keyrun (x-axis) and in the 2022 keyrun (y-axis). Each dot represent the proportion of the prey in a diet observation (year, quarter and cod (predator) size class).**

### Estimation of food ration from stomach contents data

Average daily energy consumption rates $C$ (kJ $d^{-1}$) were estimated using the cylinder gastric evacuation rate model (Andersen and Beyer, 2005a, b) by year and 1-cm predator length group for cod between 20 and 80 cm total length, amounting to 109 000 stomachs in this size range from the stomach database. Ambient temperature $T$ was assumed constant at 5°C, corresponding roughly to the average temperature experienced by cod in the Baltic Sea (Righton *et al.*, 2010). Although cod experience varying temperature throughout the year, only significant trends in average temperature regime for the cod in their preferred habitat might potentially bias our analyses. Such trends have not been shown for the Baltic Sea. We assumed constant energy densities $E_i$ for benthic prey (3.5 kJ $g^{-1}$) and consumed fishes (*Clupea harengus L.* (herring) and sprat 5.5 kJ $g^{-1}$, cod 4.0 kJ $g^{-1}$ ;Pedersen and Hislop, 2001). $E$ denotes the average energy densities (kJ $g^{-1}$) of the individually observed total stomach contents $S$ (g). Using the principle that consumption rate $C$ (kJ $d^{-1}$) on average over population and time equals evacuation rate (Pennington, 1985), and knowing cod total length $L$ (cm) and the basic evacuation rate parameter $\rho_0 = 2.43 \times 10^{-3}$, we used the parametrization of the cylinder model for cod presented in Andersen (2012):

$$C = 24\,\rho_0 L^{1.30} e^{0.083T} E^{0.15} \sqrt{S} \qquad (1)$$

In order to consider recent changes in cod consumption rate, the relationship between average quarterly consumption rate and total length (a priori parametrized as $C=aL^b$ with C the average quarterly consumption rate and L total length) was estimated separately for three different periods.

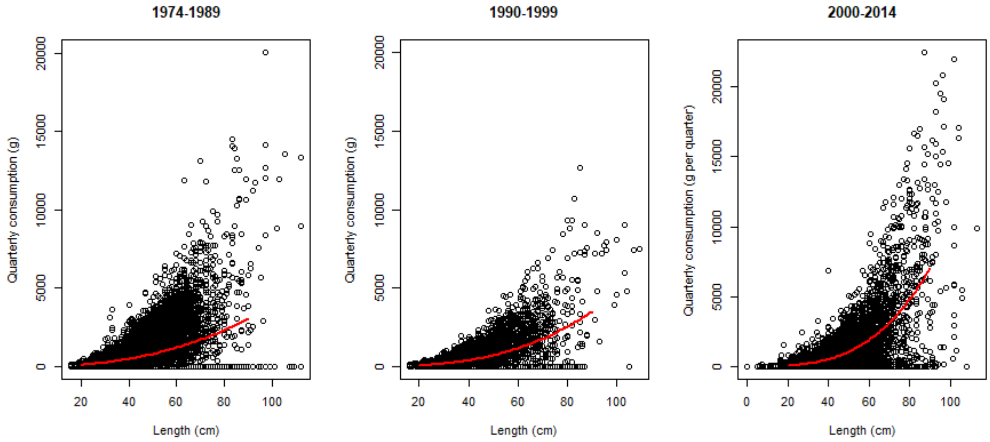


**Figure 2.3-3. Scatterplots of cod total length and estimated quarterly consumption rate. The consumption rate has been estimated separately for 1974-1989, 1990-1999 and 2000-2014 in order to account for recent changes in cod consumption (Neuenfeldt et al., 2020).**

**Table 2.3-1. Parameter estimates for the consumption rate model, C = aL$^b$.**

| Period | Parameter | Estimate | Std. Error |
|---|---|---|---|
| 1974-1989 | *a* | 0.10367 | 0.01184 |
| | *b* | 2.28617 | 0.02834 |
| 1990-1999 | *a* | 0.017408 | 0.003971 |
| | *b* | 2.713702 | 0.054565 |
| 2000-2014 | *a* | 0.003230 | 0.000354 |
| | *b* | 3.243353 | 0.025560 |

The stomach data do not include 2015-2021. For this reason, the 2000-2014 estimates were applied for 2015-2021, too.

Subsequently, average quarterly consumption was multiplied by 4 to give average yearly consumption and then distributed over quarters according to the distribution keys given in Table 2.3-2. Even though specified in the stock annex for the 2019 keyrun, the quarterly consumptions were (by an error) not redistributed over quarters, such that the same consumption was applied for all quarters.

**Table 2.3-2. Proportion of annual consumption by quarter of the year for different periods and size groups. The key was generated using all years (to account for only few data in the 3rd quarter). The length of cod (l.start and l.stop) reflect spawners and non-spawners.**

| year.start | year.stop | l.start | l.stop | q1_prop | q2_prop | q3_prop | q4_prop |
|---|---|---|---|---|---|---|---|
| 1974 | 1989 | 15 | 30 | 0.27 | 0.23 | 0.25 | 0.25 |
| 1974 | 1989 | 31 | 120 | 0.22 | 0.16 | 0.30 | 0.32 |
| 1990 | 1999 | 15 | 30 | 0.24 | 0.22 | 0.27 | 0.27 |
| 1990 | 1999 | 31 | 120 | 0.21 | 0.19 | 0.31 | 0.29 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2000 | 2019 | 15 | 30 | 0.30 | 0.16 | 0.16 | 0.38 |
| 2000 | 2019 | 31 | 120 | 0.38 | 0.19 | 0.11 | 0.32 |

### Age length keys

Age length keys (ALK) are used by SMS to transform stock number at age into stock numbers at length for the calculation of predation mortality. Length at age is derived from weight at age in the sea using a length-weight relation.  The length distribution for each age is derived from the coefficient of variation (CV) of the mean length at age as estimated from age and length observation from the BITS survey, quarter 1 and 4, 2000-2018. A year and quarter independent CV of mean length at age was derived from the estimated values by quarter (Table 2.4-1). These CV's (row "Used" in Table 2.4-1) are afterwards used to produce a length distribution around the mean length for a given age in a given year and quarter, assuming a normal distributed length distribution for each age.

**Table 2.4-1. Coefficient of variation of mean length at age derived from survey data**

```
Species         Quarter                   Age
                   0    1    2    3    4    5    6    7    8
Clupea harengus   1    NA 0.14 0.09 0.11 0.13 0.13 0.13 0.13 0.12
                  4  0.12 0.10 0.14 0.16 0.16 0.15 0.13 0.13 0.11
                Used 0.13 0.13 0.13 0.13 0.13 0.13 0.13 0.13 0.13
Gadus morhua      1    NA 0.32 0.24 0.21 0.18 0.17 0.17 0.16 0.19
                  4  0.34 0.25 0.22 0.18 0.18 0.17 0.18 0.18 0.18
                Used 0.34 0.25 0.23 0.20 0.18 0.18 0.18 0.18 0.18
Sprattus sprattus 1    NA 0.12 0.08 0.09 0.08 0.08 0.07 0.08 0.08
                  4  0.10 0.08 0.08 0.08 0.07 0.07 0.06 0.07 0.07
                Used 0.10 0.08 0.08 0.08 0.08 0.08 0.08 0.08 0.08
```

The total number of fish by length classes (Table 2.4-2) are finally calculated as the sum of contributions from each ages.  The chosen length classes depend on the length classes used in the stomach data. The "new", individual sampled stomach data (see section 2.3) have used length classes by cm and mm, however boarder length classes were used due to the low number of stomachs sampled in the individual year and quarter combinations.

The "old" pooled stomach data (see section 2.3) used larger size classes, e.g. 5-10-15 cm for sprat, in the first years of sampling. This means that the applied length classes used in the SMS configuration depends on the actual used stomach data sets used. As an example, the length classes get wider than outlined in Table 2.4-2, when both the "old" and "new" stomach data are used.  When both the "old" and "new" stomach data are used, length classes are defined for each individual year, reflecting the widest length class in the particular year.

Table 2.4-2. Default length classes used for stomach data and ALK.

| SPECIES | LOWER LENGTH (MM) | SPECIES | LOWER LENGTH (MM) | SPECIES | LOWER LENGTH (MM) |
|---|---|---|---|---|---|
| Gadus morhua | 50 | Clupea harengus | 50 | Sprattus sprattus | 50 |
| | 100 | | 70 | | 60 |
| | 150 | | 85 | | 70 |
| | 200 | | 100 | | 80 |
| | 250 | | 120 | | 90 |
| | 300 | | 140 | | 100 |
| | 350 | | 160 | | 110 |
| | 400 | | 180 | | 120 |
| | 500 | | 200 | | 130 |
| | 600 | | 220 | | 140 |
| | 700 | | 240 | | |
| | | | 260 | | |

### Predator–prey overlap

The stock area for predator cod (SD 24-32 + part of SD 23) does not completely overlap with the stock areas for herring (SD 25–29 and 32, excluding the Gulf of Riga) and sprat (SD 22-32). This will not matter, given the normally applied assumption in fisheries assessment models that individuals of the population redistributes instantaneously, such that local fishing or predation do not affect the stock distribution.

Predator–prey species overlap is a quarter dependent parameter used in the calculation of food suitability (see equation 8 in Appendix 1). By default, the spatial overlap is set to one, but it is also estimated within SMS for a few combinations, where the “quarter effect” was estimated significantly different from 1.0.

### Length–weight relations

Conversions from lengths into weights are used for some SMS configurations. The used parameters values are shown below.

**Table 2.6-1. Length (mm) weight (kg) relation for herring and sprat (W=a*l^b)**

| SPECIES | A | B |
|---|---|---|
| Herring | 2.997653e-09 | 3.136964 |
| Sprat | 3.670895e-09 | 3.107974 |

The l-w relations were estimated from BITS Q1 & Q4 data, 2000-2018 (minus 2004 data with errors). There is a statistical significant quarter effect in condition (parameter a), however this is ignored for use in SMS, until data for Quarter 2 and 3 data become available.

### References

Andersen N.G. 2012. Influences of potential predictor variables on gastric evacuation in Atlantic cod *Gadus morhua* feeding on fish prey: parameterization of a generic model. J Fish Biol 80:595–612.

Andersen N.G., Beyer J.E. 2005a. Mechanistic modelling of gastric evacuation applying the square root model to describe surface-dependent evacuation in predatory gadoids. J Fish Biol 67:1392–1412.

Andersen N.G., Beyer J.E. 2005b. Gastric evacuation of mixed stomach contents in predatory gadoids – an expanded application of the square root model to estimate food rations. Journal of Fish Biology 67:1413–1433.

Huwer B, Neuenfeldt S, Rindorf A, Andreasen H and others (2014). Study on stomach content of fish to support the assessment of good environmental status of marine food webs and the prediction of MSY after stock restoration. Final report for EU contract No. MARE/2012/02. DTU Aqua. National Institute of Aquatic Resources, Copenhagen

ICES. 1989. Report of the Study Group on cod stomach data for the Baltic. ICES CM 1989/J:2.

ICES. 1991. The international cod stomach database for the Baltic Sea and some preliminary analysis. ICES CM 1991/J:30.

ICES. 1993. Compilation of cod stomach data for the central Baltic MSVPA. ICES CM1993/J:11

ICES. 2011. Report of the Working Group on Multispecies Assessment Methods. ICES CM 2011/SSGSUE:10.

Lambert T. 1985. Gastric emptying time and assimilation efficiency in Atlantic mackerel (*Scomber scombrus*). Can J Zool 63:817–820.

Neuenfeldt, S., Bartolino, V.,Orio, A., Andersen, K. H., Andersen, N. G., Niiranen, S., Bergstro¨m, U., Ustups, D., Kulatska, N., and Casini, M. (2020). Feeding and growth of Atlantic cod (Gadus morhua L.) in the eastern Baltic Sea under environmental change. – ICES Journal of Marine Science, 77: 624–632.

Temming A, Bøhle B, Skagen DW, Knudsen FR. 2002. Gastric evacuation in mackerel: the effects of meal size, prey type and temperature. J Fish Biol 61:50–70.

# Model configuration

The configuration of the SMS model aims firstly to mimic the results from ICES single-species assessment models when SMS is run in single-species mode (no estimation of predation mortality) using the same annual M values as the single-species assessment, and secondly to configure options for estimation of predation mortality.

Appendix 2 presents the SMS configuration (option files) used for the 2022 keyrun.

## Fishing mortality

SMS uses a separable F-model (see equation 3 of the model description, in Appendix 1) while the ICES single-species assessments use XSA (and SAM as supplementary models) for herring and sprat. XSA estimates F directly from catch observations in a VPA. Further differences; SMS is using quarterly time steps while XSA is using annual time steps.

A comparison of output from the two assessments shows quite similar results for herring (Figure 3.1-1). The “SMS-single” run uses the model configuration from the 2019 keyrun, while the “SMS-single new config.” uses the suggested configuration for the 2022 keyrun. The comparison for sprat (Figure 3.1-2) show that F and SSB have the same trend, but the levels are different between the ICES version and the SMS. SSB is estimated the 1st January in SMS but at spawning time in the ICES assessment (the proportion of M and F before spawning is set to 40%) which may explain the two levels of SSB estimated. SMS estimates consistently a lower F for sprat. This is because approximately 50% of the annual yield of sprat is taken in the first quarter. Within SMS, with quarterly time steps, such catch distribution means that fishing is removing most of the individuals, before the natural mortality decreased the stock considerably. This will result in a lower F compared to the ICES-single run, with annual time steps, where F and M is assumed constant over the full year.

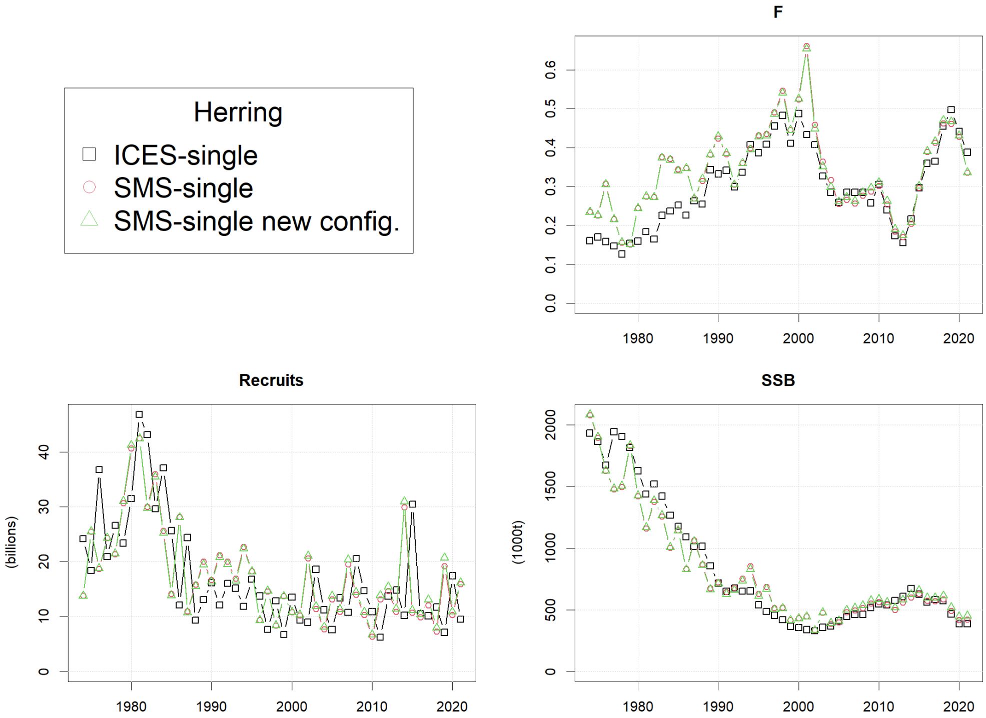


**Figure 3.1-1. Comparison of the herring assessment results from SMS assessment using fixed M (from ICES assessment) and the ICES single species XSA assessment. Recruitment is at age 0 for the SMS assessment and age 1 for the ICES assessment.**

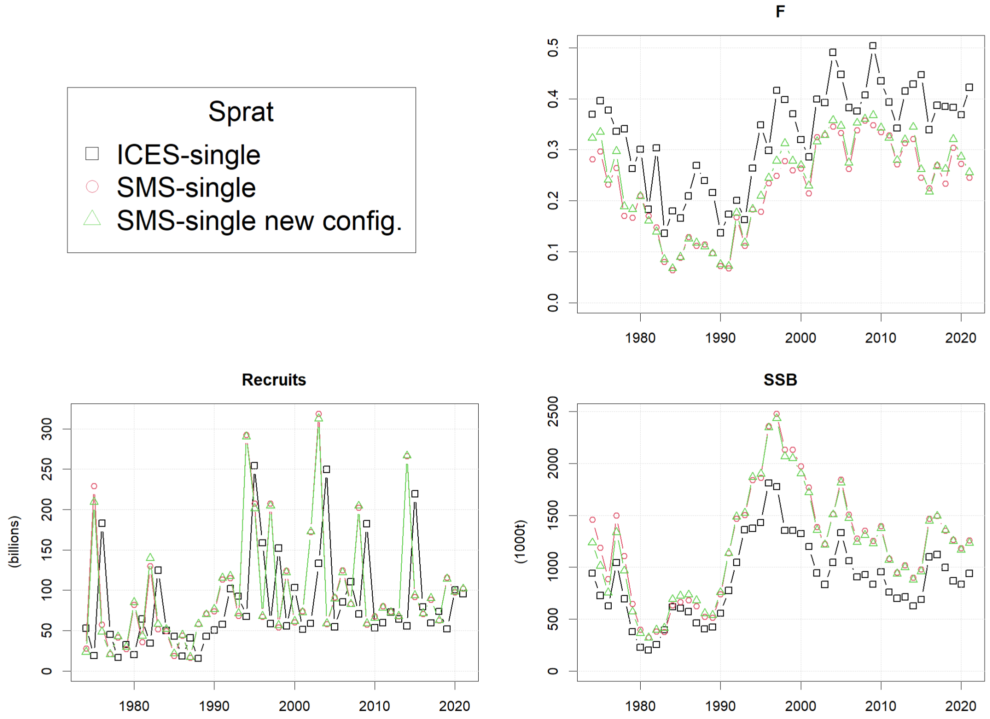


**Figure 3.1-2. Comparison of the sprat assessment results from SMS assessment using fixed M (from ICES assessment) and the ICES single species XSA assessment. Recruitment is at age 0 for the SMS assessment and age 1 for the ICES assessment.**

A closer look of the diagnostics shows that using the 2019 keyrun configuration, the catch observation variances are quite high for all ages of sprat, except for age 3. In a maximum likelihood model like SMS, there is a risk of overfitting to just one age (which might be the case for age 3), so an alternative configuration of the variance age groups (age 1, 2, 3-5, 6-7) was tried. For herring the catch observation variance for age group 2 and 3-8 seem not to be significantly different, so an alternative configuration uses the ages 1 and 2-8 groups for the variance of catch observations.

**Table 3.1-1 Observation variance of catch observations estimated from the "SMS-single" run. The configuration states that the ages 1, 2 and 3-8 have a separate variance group for herring, and ages 1, 2, 3 and 4-7 for sprat.**

| Age | Herring | Sprat |
|---|---|---|
| 1 | 0.591 | 0.761 |
| 2 | 0.384 | 0.471 |
| 3 | 0.358 | 0.384 |
| 4 | 0.358 | 0.560 |
| 5 | 0.358 | 0.560 |
| 6 | 0.358 | 0.560 |
| 7 | 0.358 | 0.560 |
| 8 | 0.358 | |

SMS estimates the catch variance by each quarter, or as done in the 2019 keyrun, with the assumption that the variances are independent of quarters, such that there is only one annual variance. As number of samples from the fishery normally follows the landings quantity, the observations variance is often highest in the periods with the lowest landings. For sprat (Figure 3.1-3) quarter 3 catches are always considerably lower than in the other quarters, which may indicate the catch variance is higher in that quarter. The effect of quarter dependent catch variances was investigated in an alternative run.

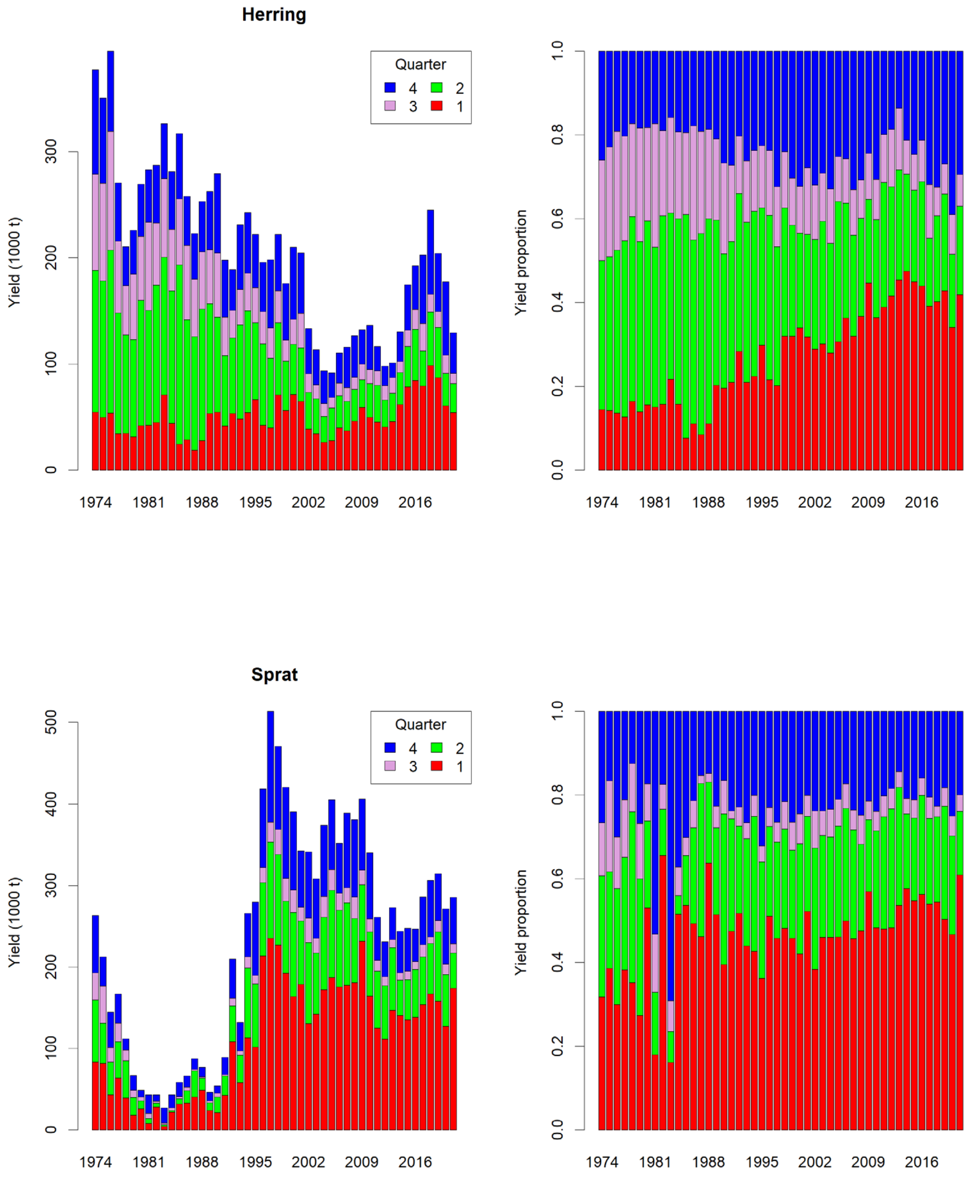


**Figure 3.1-3 Yield by year and quarter in absolute terms and as proportions for herring and sprat.**

SMS assumes that season and age effects in the model for F are constant within a specified range of years. The quarterly distribution of herring yield show a clear increasing trend of the proportion of yield taken in quarter 1, and a decreasing trend in the proportion taken in quarter 2 (Figure 3.1-3), which indicates that the F-model needs several periods with (fairly) constant season effects. The year groups used in the 2019 keyrun where 1974-1988 and 1989-2019, however these two year groups may not be sufficient to obtain a (fairly) stable selection within a period. An alternative grouping of year was tried for herring: 1974-1988, 1989-2004 and 2005-2021. For sprat, there is no clear trend in the quarterly proportion of yield, and the 2019 keyrun configuration was maintained ( 1974-1999 and 2000-2019 (2021))

#### Alternative configuration of the F-model

As suggested above, the following changes were made to the configuration of the F-model.

1. Changes of age grouping for catch observation variance for both herring and sprat
2. Change of catch observation variance from annual to seasonal for sprat
3. Change of year ranges with assumed constant selectivity for herring.

Figure 3.1-1 and Figure 3.1-2 show the effect of applying the alternative configuration, with the "SMS-single" run using the 2019 configuration and "SMS-single new config." the alternative. Sprat F is slightly higher and SSB slightly lower with the new configurations. For herring the deferens in the results is smaller.

SMS is used for estimating predations mortality, so the effects of the changes, when SMS is run with estimation of predation mortality, is presented below:

**2019 configuration of the F-model**

```
objective function (negative log likelihood):  -1423.75

Number of parameters: 304

Akaike information criterion (AIC):   -2239.51

unweighted objective function contributions:

                Catch      CPUE     SSB/R   stomach       Sum
Cod               0.0       0.0       0.0    -333.7      -334
Herring        -693.6    -146.5      -5.7       0.0      -846
Sprat          -133.6    -115.8      -7.0       0.0      -256
Sum            -827.2    -262.2     -12.7    -333.7     -1436
```

**Alternative configuration of the F-model**

```
objective function (negative log likelihood):  -1646.5

Number of parameters: 329

Akaike information criterion (AIC):   -2635

unweighted objective function contributions:

                Catch      CPUE     SSB/R   stomach       Sum
Cod               0.0       0.0       0.0    -337.6      -338
Herring        -857.6    -144.3      -6.5       0.0     -1008
Sprat          -187.9    -118.5      -7.6       0.0      -314
Sum           -1045.4    -262.7     -14.1    -337.6     -1660
```

The new configuration has a better (more negative) likelihood in total with improvements for catches for both herring and sprat. The likelihoods for survey observation (CPUE), stock recruitment (SSB/R) and stomach contents are fairly the same. The cost of the better likelihoods is the increase in model parameters from 304 to 329, however, judged from the AIC values from the two configurations, the alternative configuration is significantly better.

Values for the affected variances of catch observations by configuration are shown below:

```
2019 configuration of the F-model
sqrt(catch variance) ~ CV:
--------------------------
Herring
 1       0.587
 2       0.383
 3       0.361
 4       0.361
 5       0.361
 6       0.361
 7       0.361
 8       0.361

Sprat
 1       0.757
 2       0.475
 3       0.383
 4       0.575
 5       0.575
 6       0.575
 7       0.575
```

```
Alternative configuration of the F-model

Herring
 1       0.545
 2       0.325
 3       0.325
 4       0.325
 5       0.325
 6       0.325
 7       0.325
 8       0.325

Sprat             season
--------------------------------------
age        1       2       3       4
 1       0.646   0.750   0.957   0.619
 2       0.582   0.563   0.453   0.385
 3       0.431   0.402   0.602   0.330
 4       0.431   0.402   0.602   0.330
 5       0.431   0.402   0.602   0.330
 6       0.522   0.486   0.887   0.699
 7       0.522   0.486   0.887   0.699
```

The main effect is a much higher variance for the sprat catches in quarter 3 and lower values for the other quarters, when quarterly variances are used.

At WGSAM 2022, it was decided to use the alternative configuration of the F-model for the key-run.

### Configuring predation mortality options

The SMS model has three main options for size preferences of predators (see equations 11, 12 and 13 in the description of SMS model, Appendix 1):

1. Log normal size selection: a predator has a preferred prey size ratio and a prey twice as big as the preferred size is as attractive as another half the prey size. The preferred size ration and its variance are estimated by SMS.
2. Uniform size selection: a size preference at 1 within the range of the observed size ratio and 0 outside that ratio.
3. Constraint uniform size selection: as Uniform size selection, but the size preference ratio is constrained to exclude "outliers" in the observed size ratio.

The "Constraint uniform size selection" option was chosen for the 2012 key-run. The new stomach data available for the 2019 keyrun include more detailed data (prey length by cm group, while the old data set has prey length by 5 cm for most years) and a SMS run using the "log normal size selection" gave actually a better model fit than both the "Uniform size selection" and "Constraint uniform size selection" the (see section 5.2.3). Therefore, the "log normal size selection" and was chosen option for the 2019 and the 2022 key-runs.

### Changes in configuration/option file from the 2019 to the 2022 keyrun

To sum up: the 2022 keyrun uses a slightly different configuration of the F-model where the variance of catch observation for sprat is by quarter in the 2022 keyrun, while the 2019 keyrun used a variance for annual (summed quarterly) catches. This configuration takes the variable catch proportion between quarters, and the higher variance in quarters with a low catch, better into account. For herring, the 2022 keyrun uses three year periods to model the seasonal fishing pattern, while the 2019 keyrun uses only two year periods.

# Other issues

The SMS model, input and output for several keyruns for both the Baltic and the North Sea can be found at Github https://github.com/ices-eg/wg_WGSAM. This Github also store the SMS source code and R scripts for preparing, running and presenting results from a SMS run.

# Results of the Eastern Baltic Sea SMS keyrun

The minor changes in model configuration and input data for the 2022 keyrun compared to the 2019 version, resulted in minor changes of the output (recruitment, mean F and SSB) compared to 2019. The estimated predation mortalities (M2) from the 2022 keyrun are also similar and consistent with the M2 values from the previous keyrun.

## Keyrun summary sheet

| AREA | EASTERN BALTIC SEA |
|---|---|
| Model name | SMS |
| Type of model | Age–length structured statistical estimation model |
| Run year | 2022 |
| Predatory species | Cod |
| Prey species | Herrnig, Sprat |
| Time range | 1974–2021. |
| Time step | Quarterly |
| Area structure | Eastern Baltic Sea, ICES sub-divisions 25-29 excl Gulf of Riga |
| Stomach data | Cod: 1974-2014, ~60000 stomachs |
| Purpose of keyrun | Making historic data on natural mortality available and multispecies dynamics |
| Model changes since last keyrun in 2019 | All time-series updated (2019-20221). Diet and food rations data were recalcualted based on the same stomach observations as used in the previous keyrun. A small change in the configuration of the fishing mortality model was also made. |
| Output available at | WGSAM Sharepoint/data/EBaltic_SMS_key_run and https://github.com/ices-eg/wg_WGSAM |
| Further details in | Report of the Working Group on Multispecies Assessment Methods 2022 (WGSAM, 2022) |

### Results of the 2022 keyrun

#### Model diagnostics

The population dynamics of all species except ‘external predators’ were estimated within the model. The key-run converged and the uncertainties of parameters and key output variables were obtained from the inverse Hessian matrix. Key diagnostics (Table 5.1-1) show a reasonable fit for catch (“sqrt(catch variance) ~ CV:”) and survey indices (“sqrt(Survey variance) ~ CV:”) data. Catch and survey data fit better for herring than for sprat. The same can be seen from the catch at age residual plots (Figure 5.1-2). Herring has, in general, smaller residuals than sprat, but herring residuals show a more clustered distribution with periods of either positive or negative residuals. The survey residuals show in some cases a “year effect” with either all positive or all negative residuals within a year. This is often seen when the survey indices are based on an acoustic measurements.

The residual plot of stomach contents, Figure 5.1-4, shows a quite randomly distributed residuals for sprat. Model estimates of the stomach contents of herring seems generally higher than observed values in the period after 1990, while the opposite pattern is seen for “other food”. The same picture is seen in the boxplots of residuals (Figure 5.1-5), where the upper two rows of the plot show generally positive residuals for herring and generally negative residuals for “other food” from 1990 onwards. The bias in residuals by quarter seems limited (third row of Figure 5.1-5). The residual pattern is not independent of predator size (fourth row of Figure 5.1-5). The model overestimates the stomach contents of herring for the medium sized cod, and underestimate the stomach contents of sprat for the largest cod. This might be a result of a size dependent spatial distribution of cod.

The one step ahead residuals (Trijoulet, et al., 2023) are shown in Figure 5.1-6. They differ slightly from the standardised residual (Figure 5.1-4), but the overall pattern is the same for the two kinds of residuals.

**Table 5.1-1. SMS keyrun model diagnostics.**

```
October 09, 2022 16:36:40   run time:39 seconds

objective function (negative log likelihood):  -1646.5
Number of parameters: 329
Number of observations used in likelihood: 15714
Maximum gradient: 2.91877e-07
Akaike information criterion (AIC):   -2635
Number of observations used in the likelihood:
                            Catch    CPUE     S/R Stomach     Sum
Species: 1, Cod                 0       0       0    1599    3198
Species: 2, Herring          1536     300      48       0    3768
Species: 3, Sprat            1344     363      48       0    3510
Sum                          5760    1326     192    3198   15714


objective function weight:
                          Catch  CPUE   S/R     Stom.
Species: 1, Cod           0.00  0.00  0.00       1.00
Species: 2, Herring       1.00  1.00  0.05       0.00
Species: 3, Sprat         1.00  1.00  0.05       0.00


unweighted objective function contributions:
               Catch      CPUE     SSB/R   stomach       Sum
Cod              0.0       0.0       0.0    -337.6      -338
Herring       -857.6    -144.3      -6.5       0.0     -1008
Sprat         -187.9    -118.5      -7.6       0.0      -314
Sum          -1045.4    -262.7     -14.1    -337.6     -1660

unweighted objective function contributions (per observation):
               Catch   CPUE     S/R   Stomachs
Cod             0.00    0.00    0.00   -0.21
Herring        -0.56   -0.48   -0.14    0.00
Sprat          -0.14   -0.33   -0.16    0.00


contribution by fleet:
----------------------
Species:2, Herring
Herring Acoustic May        total: -38.824   mean:  -0.373
Herring BIAS                total:-105.430   mean:  -0.538

Species:3, Sprat
Sprat Int acoustic in Oct.  total: -89.428   mean:  -0.456
Sprat Int acoustic in May.  total: -39.596   mean:  -0.283
Sprat LAT RUS acoustic      total:  10.553   mean:   0.391


F, Year effect:
---------------
       sp. 2  sp. 3
1974:  1.000  1.000
1975:  0.976  0.992
1976:  1.318  0.691
1977:  0.933  0.792
```

```
1978:  0.681  0.483
1979:  0.659  0.458
1980:  1.081  0.507
1981:  1.208  0.366
1982:  1.180  0.306
1983:  1.592  0.181
1984:  1.550  0.145
1985:  1.426  0.192
1986:  1.478  0.272
1987:  1.133  0.259
1988:  1.355  0.253
1989:  1.000  0.237
1990:  1.111  0.204
1991:  1.004  0.210
1992:  0.798  0.555
1993:  0.939  0.381
1994:  1.031  0.628
1995:  1.120  0.730
1996:  1.120  0.865
1997:  1.265  1.004
1998:  1.401  1.133
1999:  1.143  1.010
2000:  1.362  1.000
2001:  1.705  0.857
2002:  1.163  1.177
2003:  0.899  1.238
2004:  0.755  1.340
2005:  1.000  1.279
2006:  1.042  1.012
2007:  1.003  1.289
2008:  1.087  1.321
2009:  1.126  1.327
2010:  1.175  1.229
2011:  0.995  1.180
2012:  0.727  1.038
2013:  0.665  1.173
2014:  0.808  1.280
2015:  1.180  0.981
2016:  1.551  0.814
2017:  1.652  0.996
2018:  1.853  0.972
2019:  1.828  1.190
2020:  1.662  1.057
2021:  1.262  0.951

F, season effect:
-----------------
Herring
age: 1
    1974-1988:   0.033 0.071 0.115 0.250
    1989-2004:   0.070 0.056 0.079 0.250
    2005-2021:   0.160 0.077 0.043 0.250
age: 2
    1974-1988:   0.112 0.406 0.176 0.250
    1989-2004:   0.176 0.211 0.095 0.250
    2005-2021:   0.276 0.163 0.042 0.250
age: 3 - 8
    1974-1988:   0.138 0.613 0.306 0.250
    1989-2004:   0.223 0.367 0.138 0.250
    2005-2021:   0.374 0.260 0.089 0.250

Sprat
age: 1
    1974-1999:   0.073 0.044 0.035 0.250
    2000-2021:   0.325 0.141 0.051 0.250
age: 2 - 7
    1974-1999:   0.428 0.279 0.068 0.250
    2000-2021:   0.640 0.351 0.060 0.250

F, age effect:
--------------
                0      1      2      3      4      5      6      7      8
Herring
1974-1988:  0.000  0.092  0.145  0.152  0.168  0.205  0.205  0.205  0.205
1989-2004:  0.000  0.181  0.235  0.284  0.362  0.466  0.466  0.466  0.466
2005-2021:  0.000  0.099  0.148  0.165  0.225  0.321  0.321  0.321  0.321
Sprat
```

```
1974-1999:  0.000  0.095  0.164  0.243  0.249  0.249  0.249  0.249
2000-2021:  0.000  0.108  0.142  0.181  0.196  0.196  0.196  0.196


Exploitation pattern (scaled to mean F=1)
---------------------------------------
                          0      1      2      3      4      5      6      7      8
Herring
1974-1988 season 1:       0  0.013  0.068  0.088  0.097  0.118  0.118  0.118  0.118
          season 2:       0  0.027  0.247  0.390  0.432  0.527  0.527  0.527  0.527
          season 3:   0.000  0.044  0.107  0.195  0.216  0.263  0.263  0.263  0.263
          season 4:   0.000  0.096  0.152  0.159  0.176  0.215  0.215  0.215  0.215

1989-2004 season 1:       0  0.033  0.107  0.164  0.210  0.270  0.270  0.270  0.270
          season 2:       0  0.026  0.129  0.270  0.344  0.443  0.443  0.443  0.443
          season 3:   0.000  0.037  0.058  0.101  0.129  0.167  0.167  0.167  0.167
          season 4:   0.000  0.117  0.152  0.184  0.235  0.302  0.302  0.302  0.302

2005-2021 season 1:       0  0.063  0.163  0.247  0.335  0.478  0.478  0.478  0.478
          season 2:       0  0.031  0.096  0.171  0.233  0.332  0.332  0.332  0.332
          season 3:   0.000  0.017  0.025  0.059  0.080  0.114  0.114  0.114  0.114
          season 4:   0.000  0.099  0.147  0.165  0.224  0.320  0.320  0.320  0.320

Sprat
1974-1999 season 1:       0  0.027  0.277  0.411  0.420  0.420  0.420  0.420
          season 2:       0  0.017  0.181  0.268  0.274  0.274  0.274  0.274
          season 3:   0.000  0.013  0.044  0.066  0.067  0.067  0.067  0.067
          season 4:   0.000  0.094  0.162  0.240  0.246  0.246  0.246  0.246

2000-2021 season 1:       0  0.142  0.365  0.466  0.505  0.505  0.505  0.505
          season 2:       0  0.061  0.200  0.255  0.277  0.277  0.277  0.277
          season 3:   0.000  0.022  0.034  0.044  0.047  0.047  0.047  0.047
          season 4:   0.000  0.109  0.143  0.182  0.197  0.197  0.197  0.197


sqrt(catch variance) ~ CV:
--------------------------

Herring
 1       0.545
 2       0.325
 3       0.325
 4       0.325
 5       0.325
 6       0.325
 7       0.325
 8       0.325

Sprat
              season
--------------------------------------
age        1       2       3       4

 1       0.646   0.750   0.957   0.619
 2       0.582   0.563   0.453   0.385
 3       0.431   0.402   0.602   0.330
 4       0.431   0.402   0.602   0.330
 5       0.431   0.402   0.602   0.330
 6       0.522   0.486   0.887   0.699
 7       0.522   0.486   0.887   0.699


Survey catchability:
--------------------
Herring            age 0  age 1  age 2  age 3  age 4  age 5  age 6  age 7  age 8
 Herring Acoustic May    0.430  1.054  1.797  2.503  2.503  2.503  2.503  2.503
 Herring BIAS            0.567  1.192  2.082  3.106  3.106  3.106  3.106
Sprat
 Sprat Int acous in Oct  0.469  0.706  1.022  0.945  0.945  0.945  0.945
 Sprat Int acous in May  0.303  0.699  1.054  1.024  1.024  1.024  1.024
 Sprat LAT RUS acoustic  0.329

sqrt(Survey variance) ~ CV:
---------------------------
Herring                      age 0  age 1  age 2  age 3  age 4  age 5  age 6  age 7  age 8
 Herring Acoustic May                0.38   0.38   0.38   0.44   0.44   0.44   0.44   0.44
 Herring BIAS                        0.41   0.33   0.33   0.33   0.33   0.37   0.37
Sprat
 Sprat Int acoustic in Oct.          0.47   0.35   0.35   0.38   0.38   0.38   0.38
 Sprat Int acoustic in May.          0.49   0.35   0.35   0.51   0.51   0.51   0.51
 Sprat LAT RUS acoustic              0.90
```

```
Recruit-SSB                               alfa      beta        var      sd
Herring      Geometric mean:             16.853               0.280    0.530
Sprat        Geometric mean:             18.573               0.229    0.479

Multispecies parameters
========================
stomach content variance model: Dirichlet distribution

Vulnerability pred - prey
---------------------------
           Other-food     Herring    Sprat
Cod            1.000      8.829      3.580

Size selection parameters:
---------------------------           Cod
Size selection model:               log-norm.
Sum prey sizes in likelihood:            yes
Preferred size ratio:                  5.650
Variance of size ratio:                2.674

Other food Suitability slope:
Cod                 0.3150

Stomach variance:    value    internal     max alfa0
Cod                  0.426    0.426        41.552

Predator prey season overlap
----------------------------
Predator:Cod         Other-food  Herring     Sprat
  q:1                 1           1           1
  q:2                 0.445       0.273       1
  q:3                 0.445       0.273       0.309
  q:4                 1.811       1           0.896
```

**Figure 5.1-1 Observed and model predicted catch**

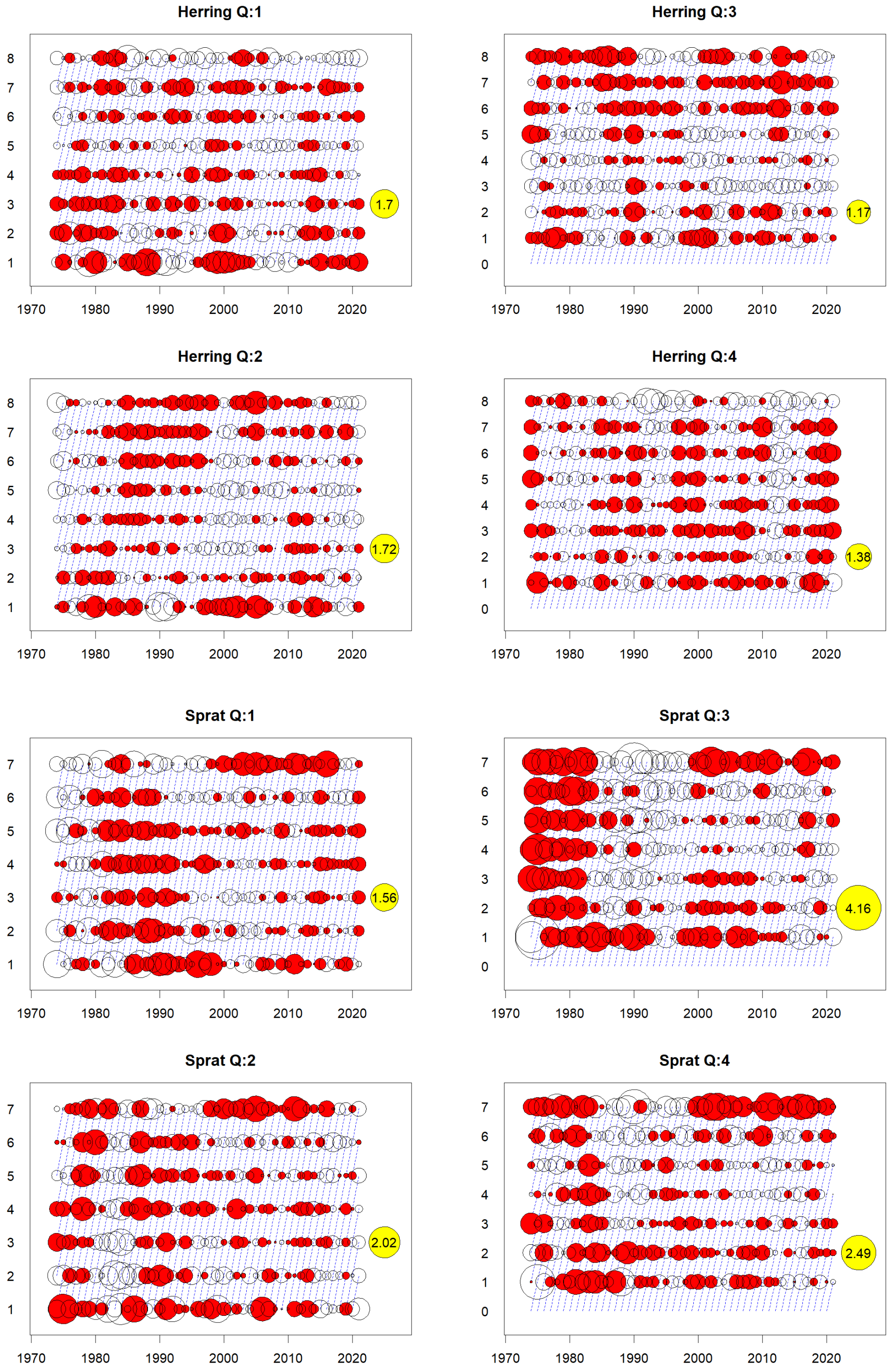


**Figure 5.1-2. Residual plots for catch-at-age observations by species and quarter. Residuals are not standardised. The red dots shows that the observed catches are larger than the model estimates. The yellow dots show the largest residual value as a reference for the dot sizes.**

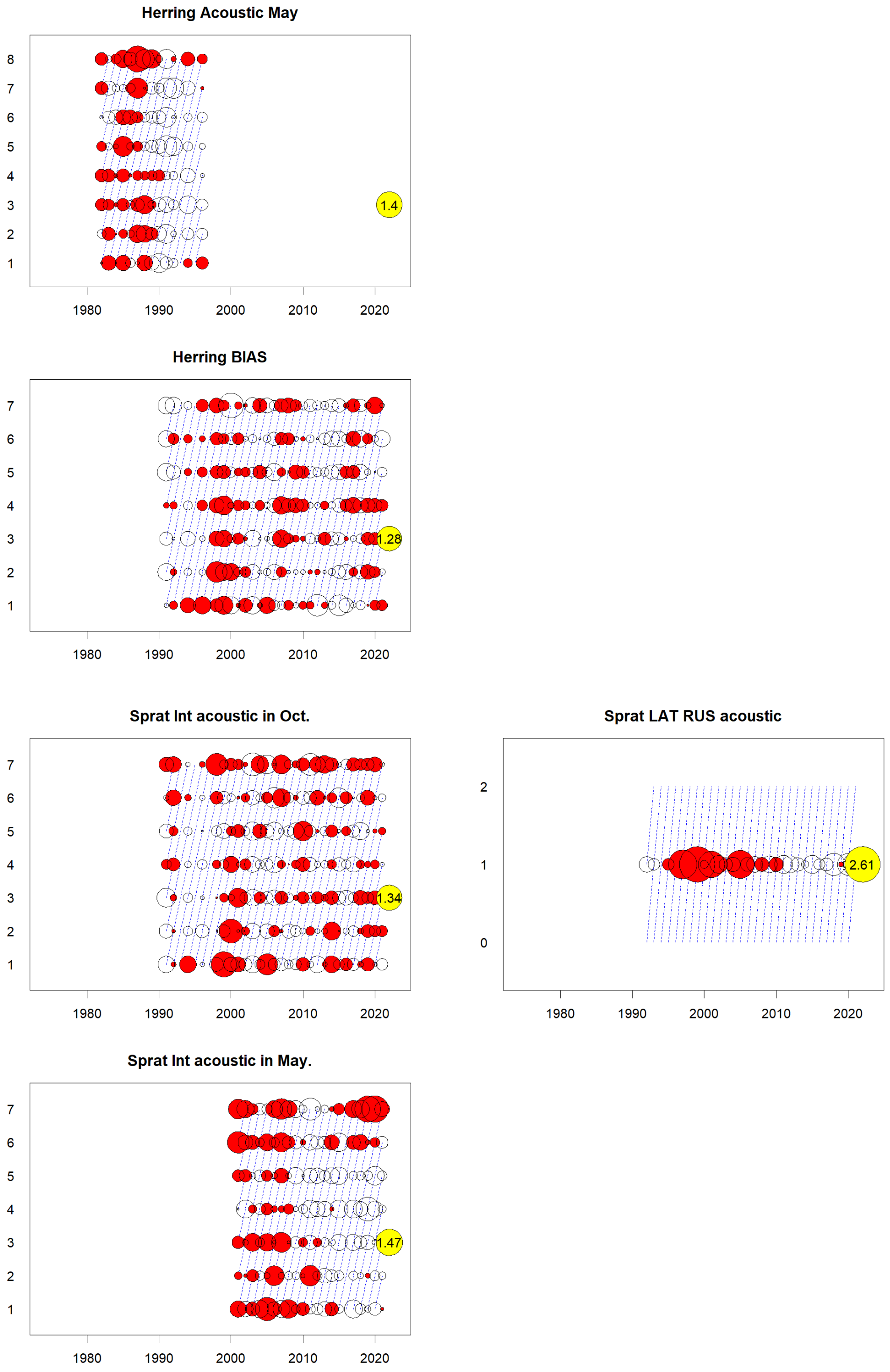


**Figure 5.1-3. Residual plots for survey, Catch per unit effort, at age observations by species and survey. Residuals are not standardised. The red dots show that the observed catches are larger than the model estimates. The yellow dots show the largest residual values per plot.**

**Figure 5.1-4. Stomach contents residuals (Standardised residuals). The y-axis show prey group and predator (cod) size class. The x-axis time period, where the upper panel is sorted by year and quarter, and lower panel sorted by quarter and year. Green dots show that the observed stomach contents are lower than the model estimates.**

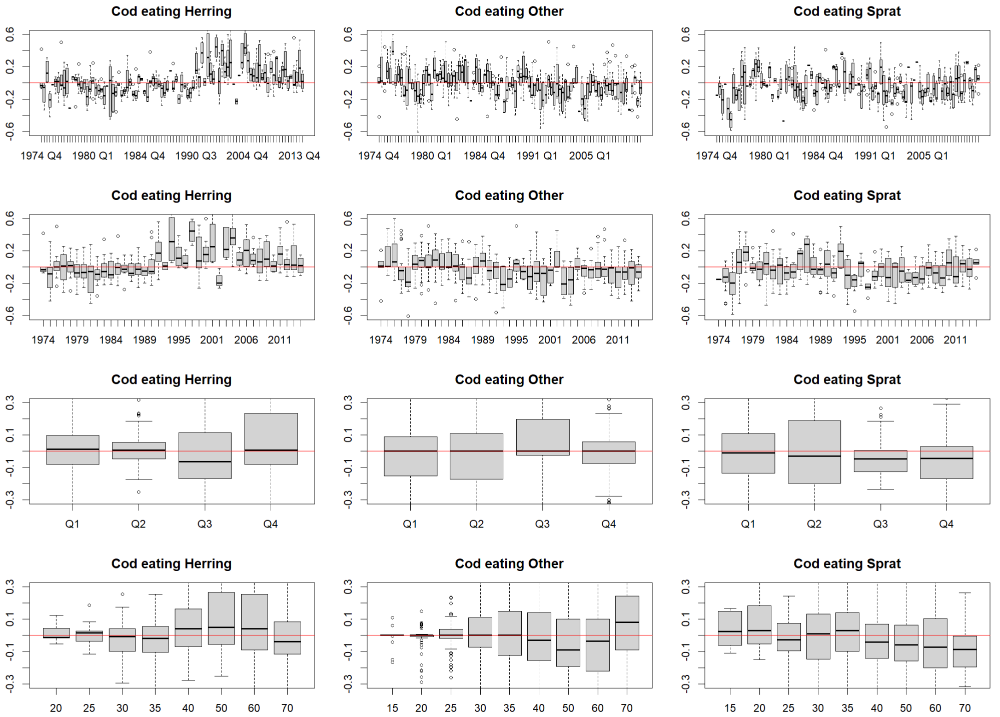


**Figure 5.1-5. Box plot of stomach contents residuals ("Raw"). Upper (first) row shows the boxplots by individual quarter and years, second row by year (quarters combined), third row by quarter and fourth row by cod size class.**

**Figure 5.1-6 Stomach contents residuals (One step ahead residuals, see Trijoulet, et al., 2023). The y-axis show prey group and predator (cod) size class. The x-axis time period, where the upper panel is sorted by year and quarter. Green dots show that the observed stomach contents are lower than the model estimates.**

References

Trijoulet, V., Albertsen, C.M., Kristensen, K., Legault C.M., Miller T.J., Nielsen, A. (2023). Model validation for compositional data in stock assessment models: Calculating residuals with correct properties. Fisheries Research, Volume 257: 106487 DOI: 10.1016/j.fishres.2022.106487

### Stock summary results

The stock summaries are presented in Figure 5.1-7 (herring) and Figure 5.1-8 (sprat).

The estimated predation mortalities (M2) are shown in details in Figure 5.1-9 and Figure 5.1-10. Total natural mortality M=M1+M2 are tabulated in Table 5.1-2 and **Error! Reference source not found.** Figure 5.1-11 shows the same data using the same scale on the y-axis and with an added smoother. The smoothed M values are tabulated in Table 5.1-4 and Table 5.1-5.

A comparison of M2 from this keyrun with M2 from the 2019 keyrun show similar values for herring (Figure 5.1-12) and for sprat (Figure 5.1-13), even though sprat M2 is now estimated higher for age 0 and 1, and herring M2 for age 0 is now higher.

Natural mortalities (M=M1+M2) estimated by SMS may be used as input to the ICES stock assessment of herring and sprat. If M values are used, WGSAM does recommend to update the full time series of M.

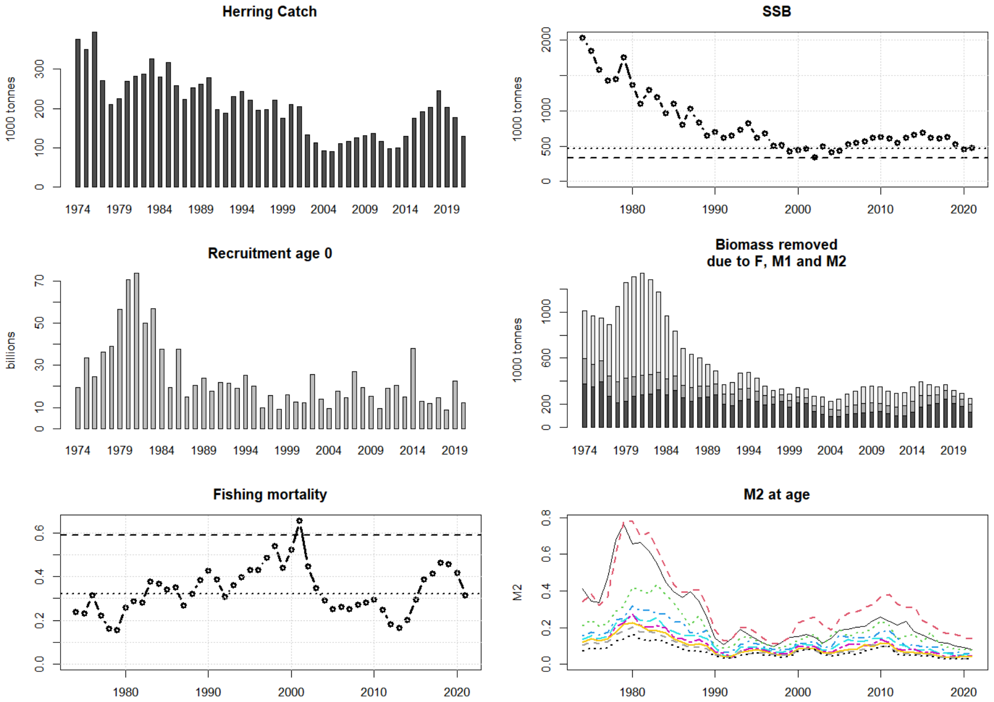


**Figure 5.1-7. SMS output for Herring. Catch weight, Recruitment, F, SSB, Biomass removed due to fishery (F), predation by SMS species (M2) and residual natural mortality (M1). The predation mortality (M2) presented by the 0-group (black solid line) is for the second half of the year. The M2 for the rest of the ages are annual values.**

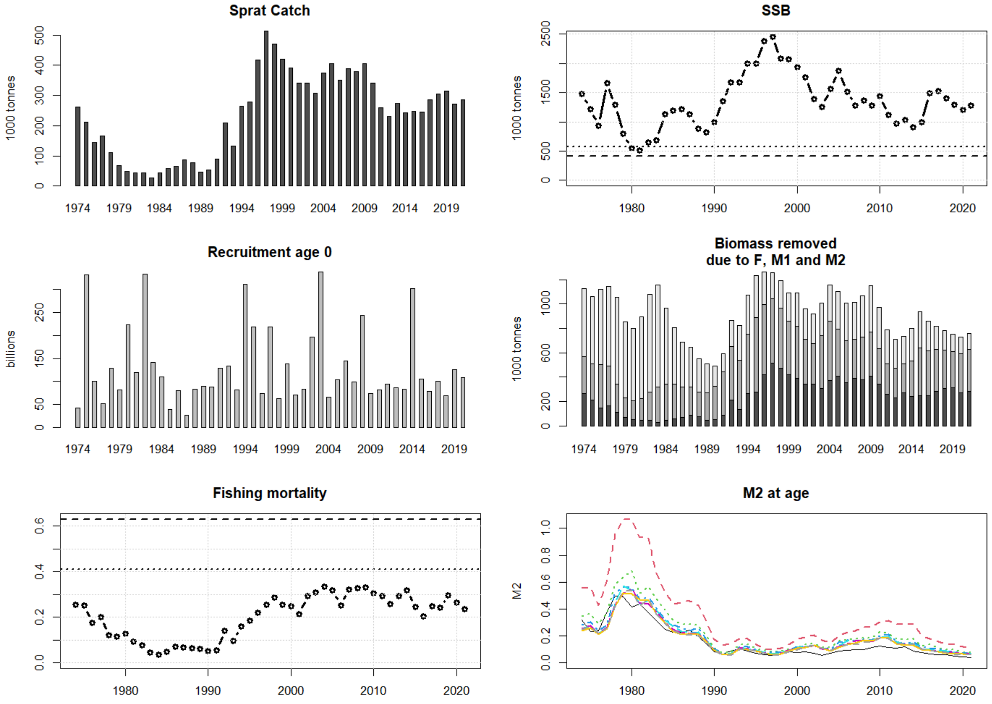


**Figure 5.1-8. SMS output for Sprat. Catch weight, Recruitment, F, SSB, Biomass removed due to fishery (F), predation by SMS species (M2) and residual natural mortality (M1). The predation mortality (M2) presented by the 0-group (black solid line) is for the second half of the year. The M2 for the rest of the ages are annual values.**

**Table 5.1-2. Herring : Natural mortality (sum of quarterly M1+M2)**

| Year/Age | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8+ |
|---|---|---|---|---|---|---|---|---|---|
| **1974** | 0.463 | 0.442 | 0.311 | 0.253 | 0.234 | 0.220 | 0.218 | 0.204 | 0.174 |
| **1975** | 0.395 | 0.486 | 0.341 | 0.277 | 0.254 | 0.239 | 0.238 | 0.223 | 0.189 |
| **1976** | 0.387 | 0.420 | 0.307 | 0.258 | 0.239 | 0.227 | 0.225 | 0.213 | 0.182 |
| **1977** | 0.527 | 0.469 | 0.328 | 0.273 | 0.253 | 0.239 | 0.237 | 0.222 | 0.189 |
| **1978** | 0.727 | 0.694 | 0.388 | 0.343 | 0.323 | 0.301 | 0.275 | 0.260 | 0.229 |
| **1979** | 0.812 | 0.874 | 0.416 | 0.353 | 0.340 | 0.326 | 0.318 | 0.283 | 0.234 |
| **1980** | 0.706 | 0.886 | 0.522 | 0.417 | 0.371 | 0.375 | 0.326 | 0.302 | 0.267 |
| **1981** | 0.716 | 0.802 | 0.503 | 0.394 | 0.340 | 0.308 | 0.308 | 0.276 | 0.236 |
| **1982** | 0.667 | 0.819 | 0.484 | 0.399 | 0.336 | 0.302 | 0.283 | 0.280 | 0.227 |
| **1983** | 0.601 | 0.730 | 0.537 | 0.379 | 0.354 | 0.312 | 0.282 | 0.265 | 0.236 |
| **1984** | 0.499 | 0.626 | 0.480 | 0.372 | 0.300 | 0.299 | 0.271 | 0.246 | 0.220 |
| **1985** | 0.447 | 0.537 | 0.436 | 0.332 | 0.279 | 0.248 | 0.241 | 0.227 | 0.205 |
| **1986** | 0.415 | 0.491 | 0.374 | 0.331 | 0.260 | 0.239 | 0.220 | 0.207 | 0.183 |
| **1987** | 0.448 | 0.500 | 0.312 | 0.271 | 0.254 | 0.219 | 0.203 | 0.192 | 0.172 |
| **1988** | 0.394 | 0.508 | 0.368 | 0.266 | 0.253 | 0.235 | 0.212 | 0.194 | 0.173 |
| **1989** | 0.303 | 0.423 | 0.285 | 0.287 | 0.239 | 0.212 | 0.201 | 0.183 | 0.165 |
| **1990** | 0.195 | 0.285 | 0.205 | 0.185 | 0.193 | 0.167 | 0.160 | 0.154 | 0.147 |
| **1991** | 0.159 | 0.227 | 0.189 | 0.164 | 0.148 | 0.158 | 0.140 | 0.144 | 0.134 |
| **1992** | 0.190 | 0.236 | 0.192 | 0.170 | 0.145 | 0.137 | 0.146 | 0.134 | 0.131 |
| **1993** | 0.240 | 0.296 | 0.245 | 0.207 | 0.191 | 0.173 | 0.163 | 0.170 | 0.149 |
| **1994** | 0.211 | 0.304 | 0.251 | 0.221 | 0.193 | 0.184 | 0.172 | 0.157 | 0.156 |
| **1995** | 0.185 | 0.268 | 0.228 | 0.211 | 0.194 | 0.183 | 0.178 | 0.166 | 0.163 |
| **1996** | 0.161 | 0.235 | 0.213 | 0.192 | 0.183 | 0.175 | 0.167 | 0.162 | 0.150 |
| **1997** | 0.146 | 0.212 | 0.196 | 0.178 | 0.170 | 0.161 | 0.156 | 0.152 | 0.147 |
| **1998** | 0.171 | 0.217 | 0.189 | 0.176 | 0.162 | 0.155 | 0.148 | 0.146 | 0.135 |
| **1999** | 0.195 | 0.245 | 0.206 | 0.184 | 0.175 | 0.163 | 0.153 | 0.150 | 0.139 |
| **2000** | 0.202 | 0.324 | 0.233 | 0.219 | 0.207 | 0.198 | 0.183 | 0.174 | 0.172 |
| **2001** | 0.211 | 0.343 | 0.246 | 0.213 | 0.208 | 0.193 | 0.186 | 0.182 | 0.179 |
| **2002** | 0.199 | 0.358 | 0.263 | 0.226 | 0.203 | 0.195 | 0.184 | 0.177 | 0.181 |
| **2003** | 0.165 | 0.313 | 0.208 | 0.189 | 0.178 | 0.172 | 0.164 | 0.156 | 0.153 |
| **2004** | 0.193 | 0.286 | 0.259 | 0.197 | 0.183 | 0.165 | 0.159 | 0.153 | 0.146 |
| **2005** | 0.237 | 0.345 | 0.290 | 0.259 | 0.211 | 0.187 | 0.171 | 0.163 | 0.153 |
| **2006** | 0.239 | 0.378 | 0.252 | 0.249 | 0.234 | 0.209 | 0.180 | 0.170 | 0.161 |
| **2007** | 0.252 | 0.387 | 0.270 | 0.243 | 0.222 | 0.210 | 0.181 | 0.169 | 0.154 |
| **2008** | 0.257 | 0.410 | 0.272 | 0.248 | 0.227 | 0.201 | 0.209 | 0.183 | 0.171 |
| **2009** | 0.287 | 0.425 | 0.299 | 0.248 | 0.239 | 0.211 | 0.194 | 0.209 | 0.185 |
| **2010** | 0.307 | 0.470 | 0.340 | 0.277 | 0.240 | 0.234 | 0.215 | 0.200 | 0.197 |
| **2011** | 0.286 | 0.479 | 0.309 | 0.283 | 0.242 | 0.219 | 0.213 | 0.194 | 0.195 |
| **2012** | 0.266 | 0.428 | 0.236 | 0.242 | 0.210 | 0.184 | 0.175 | 0.164 | 0.154 |
| **2013** | 0.282 | 0.409 | 0.256 | 0.196 | 0.201 | 0.177 | 0.163 | 0.157 | 0.150 |
| **2014** | 0.227 | 0.413 | 0.262 | 0.209 | 0.176 | 0.182 | 0.162 | 0.155 | 0.149 |
| **2015** | 0.210 | 0.338 | 0.221 | 0.194 | 0.175 | 0.160 | 0.159 | 0.150 | 0.145 |
| **2016** | 0.188 | 0.317 | 0.276 | 0.191 | 0.179 | 0.167 | 0.157 | 0.153 | 0.146 |
| **2017** | 0.169 | 0.284 | 0.213 | 0.196 | 0.162 | 0.155 | 0.146 | 0.137 | 0.134 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **2018** | 0.166 | 0.266 | 0.193 | 0.161 | 0.159 | 0.138 | 0.137 | 0.135 | 0.128 |
| **2019** | 0.148 | 0.262 | 0.181 | 0.167 | 0.150 | 0.150 | 0.136 | 0.133 | 0.129 |
| **2020** | 0.143 | 0.246 | 0.178 | 0.158 | 0.151 | 0.141 | 0.142 | 0.130 | 0.129 |
| **2021** | 0.130 | 0.240 | 0.176 | 0.164 | 0.151 | 0.147 | 0.140 | 0.138 | 0.132 |

Table 5.1-3. Sprat : Natural mortality (sum of quarterly M1+M2)

| **Year/Age** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** |
|---|---|---|---|---|---|---|---|---|
| **1974** | 0.416 | 0.755 | 0.543 | 0.478 | 0.452 | 0.452 | 0.439 | 0.449 |
| **1975** | 0.331 | 0.762 | 0.566 | 0.503 | 0.476 | 0.476 | 0.462 | 0.472 |
| **1976** | 0.336 | 0.623 | 0.485 | 0.437 | 0.418 | 0.418 | 0.409 | 0.416 |
| **1977** | 0.482 | 0.816 | 0.572 | 0.485 | 0.464 | 0.464 | 0.451 | 0.460 |
| **1978** | 0.594 | 1.158 | 0.784 | 0.720 | 0.640 | 0.629 | 0.619 | 0.619 |
| **1979** | 0.591 | 1.269 | 0.835 | 0.766 | 0.771 | 0.713 | 0.718 | 0.733 |
| **1980** | 0.516 | 1.264 | 0.886 | 0.757 | 0.741 | 0.751 | 0.713 | 0.731 |
| **1981** | 0.535 | 1.130 | 0.717 | 0.676 | 0.638 | 0.641 | 0.668 | 0.619 |
| **1982** | 0.473 | 1.124 | 0.768 | 0.684 | 0.665 | 0.637 | 0.666 | 0.674 |
| **1983** | 0.413 | 0.867 | 0.676 | 0.608 | 0.590 | 0.576 | 0.564 | 0.561 |
| **1984** | 0.349 | 0.722 | 0.595 | 0.522 | 0.517 | 0.501 | 0.495 | 0.493 |
| **1985** | 0.327 | 0.636 | 0.517 | 0.483 | 0.468 | 0.450 | 0.434 | 0.443 |
| **1986** | 0.327 | 0.651 | 0.486 | 0.461 | 0.434 | 0.419 | 0.413 | 0.410 |
| **1987** | 0.343 | 0.656 | 0.485 | 0.439 | 0.421 | 0.416 | 0.416 | 0.405 |
| **1988** | 0.297 | 0.626 | 0.476 | 0.461 | 0.430 | 0.414 | 0.411 | 0.400 |
| **1989** | 0.242 | 0.515 | 0.404 | 0.375 | 0.369 | 0.361 | 0.358 | 0.354 |
| **1990** | 0.180 | 0.371 | 0.308 | 0.303 | 0.297 | 0.291 | 0.293 | 0.286 |
| **1991** | 0.166 | 0.330 | 0.270 | 0.267 | 0.262 | 0.260 | 0.259 | 0.260 |
| **1992** | 0.187 | 0.346 | 0.280 | 0.268 | 0.264 | 0.261 | 0.259 | 0.257 |
| **1993** | 0.206 | 0.380 | 0.338 | 0.322 | 0.312 | 0.308 | 0.304 | 0.298 |
| **1994** | 0.185 | 0.378 | 0.334 | 0.317 | 0.305 | 0.303 | 0.299 | 0.298 |
| **1995** | 0.169 | 0.334 | 0.301 | 0.299 | 0.292 | 0.287 | 0.286 | 0.285 |
| **1996** | 0.158 | 0.305 | 0.293 | 0.279 | 0.277 | 0.271 | 0.270 | 0.270 |
| **1997** | 0.153 | 0.298 | 0.280 | 0.274 | 0.266 | 0.259 | 0.258 | 0.256 |
| **1998** | 0.168 | 0.307 | 0.286 | 0.280 | 0.277 | 0.269 | 0.267 | 0.268 |
| **1999** | 0.180 | 0.337 | 0.304 | 0.292 | 0.293 | 0.291 | 0.284 | 0.281 |
| **2000** | 0.175 | 0.376 | 0.316 | 0.318 | 0.313 | 0.309 | 0.306 | 0.298 |
| **2001** | 0.179 | 0.391 | 0.333 | 0.319 | 0.320 | 0.314 | 0.316 | 0.317 |
| **2002** | 0.171 | 0.405 | 0.341 | 0.337 | 0.330 | 0.330 | 0.329 | 0.330 |
| **2003** | 0.154 | 0.366 | 0.315 | 0.309 | 0.308 | 0.303 | 0.307 | 0.308 |
| **2004** | 0.169 | 0.345 | 0.316 | 0.296 | 0.289 | 0.292 | 0.290 | 0.291 |
| **2005** | 0.189 | 0.399 | 0.363 | 0.349 | 0.326 | 0.321 | 0.316 | 0.321 |
| **2006** | 0.190 | 0.429 | 0.375 | 0.369 | 0.360 | 0.341 | 0.335 | 0.336 |
| **2007** | 0.197 | 0.440 | 0.380 | 0.362 | 0.362 | 0.360 | 0.346 | 0.335 |
| **2008** | 0.199 | 0.466 | 0.382 | 0.373 | 0.363 | 0.367 | 0.369 | 0.350 |
| **2009** | 0.213 | 0.465 | 0.383 | 0.368 | 0.361 | 0.358 | 0.363 | 0.356 |
| **2010** | 0.222 | 0.504 | 0.430 | 0.401 | 0.390 | 0.387 | 0.383 | 0.386 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **2011** | 0.213 | 0.515 | 0.417 | 0.409 | 0.394 | 0.381 | 0.383 | 0.377 |
| **2012** | 0.207 | 0.487 | 0.380 | 0.357 | 0.356 | 0.347 | 0.343 | 0.345 |
| **2013** | 0.217 | 0.488 | 0.372 | 0.343 | 0.335 | 0.334 | 0.332 | 0.333 |
| **2014** | 0.188 | 0.491 | 0.378 | 0.356 | 0.338 | 0.332 | 0.333 | 0.338 |
| **2015** | 0.178 | 0.400 | 0.327 | 0.314 | 0.306 | 0.301 | 0.297 | 0.303 |
| **2016** | 0.167 | 0.376 | 0.336 | 0.309 | 0.298 | 0.295 | 0.291 | 0.292 |
| **2017** | 0.159 | 0.355 | 0.309 | 0.301 | 0.286 | 0.280 | 0.280 | 0.280 |
| **2018** | 0.159 | 0.342 | 0.296 | 0.288 | 0.285 | 0.276 | 0.271 | 0.274 |
| **2019** | 0.148 | 0.336 | 0.297 | 0.284 | 0.279 | 0.279 | 0.267 | 0.267 |
| **2020** | 0.146 | 0.321 | 0.277 | 0.273 | 0.268 | 0.265 | 0.266 | 0.260 |
| **2021** | 0.139 | 0.310 | 0.276 | 0.269 | 0.266 | 0.263 | 0.258 | 0.262 |

**Table 5.1-4. Herring GAM-Smoothed Natural mortality (sum of quarterly M1 (=0.1)+M2)**

| Year/Age | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| **1974** | 0.362 | 0.361 | 0.282 | 0.234 | 0.218 | 0.205 | 0.206 | 0.193 | 0.166 |
| **1975** | 0.449 | 0.460 | 0.316 | 0.263 | 0.245 | 0.232 | 0.228 | 0.214 | 0.182 |
| **1976** | 0.530 | 0.554 | 0.348 | 0.291 | 0.271 | 0.257 | 0.250 | 0.233 | 0.198 |
| **1977** | 0.601 | 0.637 | 0.380 | 0.317 | 0.294 | 0.280 | 0.269 | 0.250 | 0.213 |
| **1978** | 0.656 | 0.706 | 0.409 | 0.340 | 0.313 | 0.298 | 0.284 | 0.263 | 0.224 |
| **1979** | 0.691 | 0.753 | 0.436 | 0.359 | 0.328 | 0.311 | 0.294 | 0.273 | 0.233 |
| **1980** | 0.700 | 0.776 | 0.460 | 0.374 | 0.336 | 0.318 | 0.299 | 0.277 | 0.237 |
| **1981** | 0.686 | 0.775 | 0.478 | 0.383 | 0.339 | 0.318 | 0.298 | 0.276 | 0.237 |
| **1982** | 0.653 | 0.755 | 0.489 | 0.385 | 0.336 | 0.312 | 0.291 | 0.271 | 0.234 |
| **1983** | 0.607 | 0.718 | 0.489 | 0.381 | 0.328 | 0.302 | 0.281 | 0.261 | 0.227 |
| **1984** | 0.553 | 0.670 | 0.476 | 0.369 | 0.315 | 0.288 | 0.267 | 0.249 | 0.218 |
| **1985** | 0.496 | 0.614 | 0.448 | 0.349 | 0.297 | 0.270 | 0.250 | 0.234 | 0.206 |
| **1986** | 0.440 | 0.553 | 0.408 | 0.323 | 0.276 | 0.250 | 0.232 | 0.217 | 0.193 |
| **1987** | 0.385 | 0.491 | 0.360 | 0.292 | 0.254 | 0.230 | 0.213 | 0.200 | 0.180 |
| **1988** | 0.334 | 0.432 | 0.311 | 0.262 | 0.232 | 0.211 | 0.196 | 0.184 | 0.167 |
| **1989** | 0.290 | 0.377 | 0.268 | 0.235 | 0.212 | 0.194 | 0.181 | 0.170 | 0.157 |
| **1990** | 0.253 | 0.332 | 0.236 | 0.214 | 0.197 | 0.181 | 0.169 | 0.160 | 0.149 |
| **1991** | 0.225 | 0.297 | 0.218 | 0.201 | 0.186 | 0.172 | 0.162 | 0.154 | 0.144 |
| **1992** | 0.204 | 0.271 | 0.210 | 0.194 | 0.179 | 0.168 | 0.159 | 0.152 | 0.142 |
| **1993** | 0.190 | 0.254 | 0.210 | 0.192 | 0.176 | 0.166 | 0.158 | 0.152 | 0.143 |
| **1994** | 0.181 | 0.245 | 0.214 | 0.192 | 0.175 | 0.166 | 0.159 | 0.153 | 0.144 |
| **1995** | 0.177 | 0.243 | 0.218 | 0.193 | 0.176 | 0.167 | 0.161 | 0.156 | 0.147 |
| **1996** | 0.175 | 0.246 | 0.219 | 0.193 | 0.177 | 0.169 | 0.163 | 0.158 | 0.150 |
| **1997** | 0.175 | 0.253 | 0.218 | 0.193 | 0.179 | 0.171 | 0.164 | 0.160 | 0.153 |
| **1998** | 0.178 | 0.263 | 0.218 | 0.194 | 0.181 | 0.173 | 0.166 | 0.162 | 0.156 |
| **1999** | 0.181 | 0.275 | 0.218 | 0.195 | 0.184 | 0.175 | 0.167 | 0.163 | 0.158 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **2000** | 0.186 | 0.287 | 0.220 | 0.198 | 0.187 | 0.177 | 0.169 | 0.164 | 0.160 |
| **2001** | 0.190 | 0.299 | 0.225 | 0.203 | 0.191 | 0.180 | 0.170 | 0.164 | 0.160 |
| **2002** | 0.195 | 0.310 | 0.233 | 0.209 | 0.195 | 0.182 | 0.171 | 0.164 | 0.159 |
| **2003** | 0.202 | 0.322 | 0.243 | 0.216 | 0.200 | 0.185 | 0.172 | 0.164 | 0.158 |
| **2004** | 0.210 | 0.336 | 0.254 | 0.224 | 0.205 | 0.189 | 0.174 | 0.166 | 0.158 |
| **2005** | 0.222 | 0.353 | 0.265 | 0.233 | 0.211 | 0.193 | 0.178 | 0.168 | 0.160 |
| **2006** | 0.236 | 0.374 | 0.275 | 0.242 | 0.218 | 0.198 | 0.182 | 0.172 | 0.163 |
| **2007** | 0.252 | 0.397 | 0.283 | 0.249 | 0.224 | 0.203 | 0.188 | 0.178 | 0.168 |
| **2008** | 0.268 | 0.419 | 0.289 | 0.255 | 0.229 | 0.208 | 0.194 | 0.183 | 0.173 |
| **2009** | 0.281 | 0.436 | 0.292 | 0.258 | 0.231 | 0.211 | 0.197 | 0.187 | 0.177 |
| **2010** | 0.287 | 0.446 | 0.292 | 0.258 | 0.231 | 0.211 | 0.198 | 0.188 | 0.178 |
| **2011** | 0.285 | 0.444 | 0.288 | 0.252 | 0.226 | 0.207 | 0.195 | 0.186 | 0.176 |
| **2012** | 0.275 | 0.431 | 0.280 | 0.242 | 0.217 | 0.199 | 0.188 | 0.179 | 0.171 |
| **2013** | 0.258 | 0.410 | 0.269 | 0.229 | 0.205 | 0.190 | 0.179 | 0.170 | 0.163 |
| **2014** | 0.237 | 0.383 | 0.256 | 0.214 | 0.193 | 0.179 | 0.168 | 0.160 | 0.154 |
| **2015** | 0.215 | 0.354 | 0.243 | 0.200 | 0.180 | 0.168 | 0.158 | 0.150 | 0.145 |
| **2016** | 0.195 | 0.327 | 0.229 | 0.188 | 0.170 | 0.159 | 0.149 | 0.142 | 0.138 |
| **2017** | 0.178 | 0.302 | 0.216 | 0.178 | 0.162 | 0.153 | 0.144 | 0.137 | 0.133 |
| **2018** | 0.163 | 0.281 | 0.204 | 0.171 | 0.157 | 0.148 | 0.141 | 0.134 | 0.131 |
| **2019** | 0.150 | 0.262 | 0.192 | 0.166 | 0.153 | 0.146 | 0.139 | 0.134 | 0.130 |
| **2020** | 0.139 | 0.244 | 0.180 | 0.163 | 0.151 | 0.144 | 0.139 | 0.134 | 0.130 |
| **2021** | 0.128 | 0.226 | 0.169 | 0.160 | 0.149 | 0.143 | 0.139 | 0.135 | 0.130 |

**Table 5.1-5. Sprat GAM-Smoothed Natural mortality (sum of quarterly M1 (=0.1)+M2).**

| **Year/Age** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** |
|---|---|---|---|---|---|---|---|---|
| **1974** | 0.340 | 0.605 | 0.476 | 0.420 | 0.392 | 0.393 | 0.375 | 0.389 |
| **1975** | 0.394 | 0.750 | 0.552 | 0.489 | 0.461 | 0.459 | 0.445 | 0.456 |
| **1976** | 0.444 | 0.885 | 0.623 | 0.554 | 0.526 | 0.521 | 0.511 | 0.518 |
| **1977** | 0.487 | 1.000 | 0.685 | 0.611 | 0.583 | 0.575 | 0.568 | 0.573 |
| **1978** | 0.517 | 1.086 | 0.733 | 0.655 | 0.627 | 0.617 | 0.613 | 0.615 |
| **1979** | 0.532 | 1.133 | 0.762 | 0.682 | 0.655 | 0.643 | 0.641 | 0.641 |
| **1980** | 0.528 | 1.133 | 0.768 | 0.689 | 0.663 | 0.650 | 0.648 | 0.647 |
| **1981** | 0.508 | 1.091 | 0.753 | 0.677 | 0.653 | 0.638 | 0.637 | 0.635 |
| **1982** | 0.475 | 1.020 | 0.722 | 0.651 | 0.628 | 0.612 | 0.611 | 0.608 |
| **1983** | 0.436 | 0.930 | 0.678 | 0.613 | 0.593 | 0.577 | 0.575 | 0.572 |
| **1984** | 0.394 | 0.835 | 0.627 | 0.570 | 0.551 | 0.535 | 0.533 | 0.529 |
| **1985** | 0.355 | 0.744 | 0.573 | 0.524 | 0.507 | 0.492 | 0.489 | 0.485 |
| **1986** | 0.320 | 0.664 | 0.519 | 0.478 | 0.462 | 0.449 | 0.446 | 0.442 |
| **1987** | 0.290 | 0.595 | 0.467 | 0.435 | 0.420 | 0.408 | 0.405 | 0.401 |
| **1988** | 0.263 | 0.534 | 0.419 | 0.395 | 0.381 | 0.371 | 0.368 | 0.363 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **1989** | 0.240 | 0.481 | 0.378 | 0.360 | 0.347 | 0.339 | 0.336 | 0.332 |
| **1990** | 0.220 | 0.436 | 0.346 | 0.333 | 0.321 | 0.314 | 0.311 | 0.307 |
| **1991** | 0.204 | 0.397 | 0.323 | 0.313 | 0.302 | 0.296 | 0.294 | 0.290 |
| **1992** | 0.190 | 0.367 | 0.309 | 0.299 | 0.290 | 0.285 | 0.282 | 0.279 |
| **1993** | 0.180 | 0.343 | 0.300 | 0.291 | 0.283 | 0.278 | 0.276 | 0.273 |
| **1994** | 0.172 | 0.327 | 0.296 | 0.287 | 0.280 | 0.276 | 0.273 | 0.271 |
| **1995** | 0.167 | 0.320 | 0.295 | 0.286 | 0.280 | 0.276 | 0.273 | 0.272 |
| **1996** | 0.165 | 0.320 | 0.296 | 0.286 | 0.282 | 0.278 | 0.276 | 0.275 |
| **1997** | 0.165 | 0.326 | 0.297 | 0.289 | 0.285 | 0.281 | 0.280 | 0.278 |
| **1998** | 0.166 | 0.335 | 0.300 | 0.292 | 0.289 | 0.285 | 0.284 | 0.283 |
| **1999** | 0.168 | 0.346 | 0.304 | 0.297 | 0.294 | 0.290 | 0.289 | 0.288 |
| **2000** | 0.170 | 0.356 | 0.310 | 0.303 | 0.300 | 0.296 | 0.295 | 0.294 |
| **2001** | 0.171 | 0.364 | 0.316 | 0.309 | 0.305 | 0.301 | 0.300 | 0.299 |
| **2002** | 0.172 | 0.370 | 0.324 | 0.316 | 0.311 | 0.307 | 0.305 | 0.304 |
| **2003** | 0.174 | 0.377 | 0.333 | 0.324 | 0.318 | 0.313 | 0.311 | 0.309 |
| **2004** | 0.176 | 0.386 | 0.343 | 0.333 | 0.325 | 0.321 | 0.317 | 0.316 |
| **2005** | 0.181 | 0.399 | 0.354 | 0.343 | 0.334 | 0.329 | 0.326 | 0.324 |
| **2006** | 0.187 | 0.419 | 0.366 | 0.353 | 0.344 | 0.340 | 0.336 | 0.333 |
| **2007** | 0.196 | 0.442 | 0.379 | 0.364 | 0.355 | 0.350 | 0.348 | 0.344 |
| **2008** | 0.204 | 0.466 | 0.390 | 0.374 | 0.365 | 0.360 | 0.358 | 0.354 |
| **2009** | 0.211 | 0.486 | 0.398 | 0.381 | 0.372 | 0.367 | 0.366 | 0.361 |
| **2010** | 0.215 | 0.499 | 0.402 | 0.383 | 0.374 | 0.369 | 0.369 | 0.364 |
| **2011** | 0.215 | 0.501 | 0.400 | 0.380 | 0.371 | 0.366 | 0.365 | 0.361 |
| **2012** | 0.210 | 0.491 | 0.391 | 0.371 | 0.361 | 0.356 | 0.355 | 0.353 |
| **2013** | 0.202 | 0.472 | 0.378 | 0.358 | 0.348 | 0.342 | 0.341 | 0.340 |
| **2014** | 0.192 | 0.447 | 0.362 | 0.343 | 0.332 | 0.326 | 0.324 | 0.325 |
| **2015** | 0.182 | 0.420 | 0.346 | 0.327 | 0.316 | 0.310 | 0.307 | 0.310 |
| **2016** | 0.171 | 0.394 | 0.330 | 0.312 | 0.302 | 0.296 | 0.293 | 0.296 |
| **2017** | 0.163 | 0.372 | 0.316 | 0.300 | 0.291 | 0.285 | 0.282 | 0.285 |
| **2018** | 0.156 | 0.352 | 0.303 | 0.290 | 0.282 | 0.277 | 0.273 | 0.276 |
| **2019** | 0.149 | 0.334 | 0.292 | 0.281 | 0.275 | 0.271 | 0.267 | 0.268 |
| **2020** | 0.144 | 0.317 | 0.281 | 0.274 | 0.269 | 0.266 | 0.262 | 0.262 |
| **2021** | 0.139 | 0.301 | 0.271 | 0.267 | 0.264 | 0.262 | 0.258 | 0.256 |

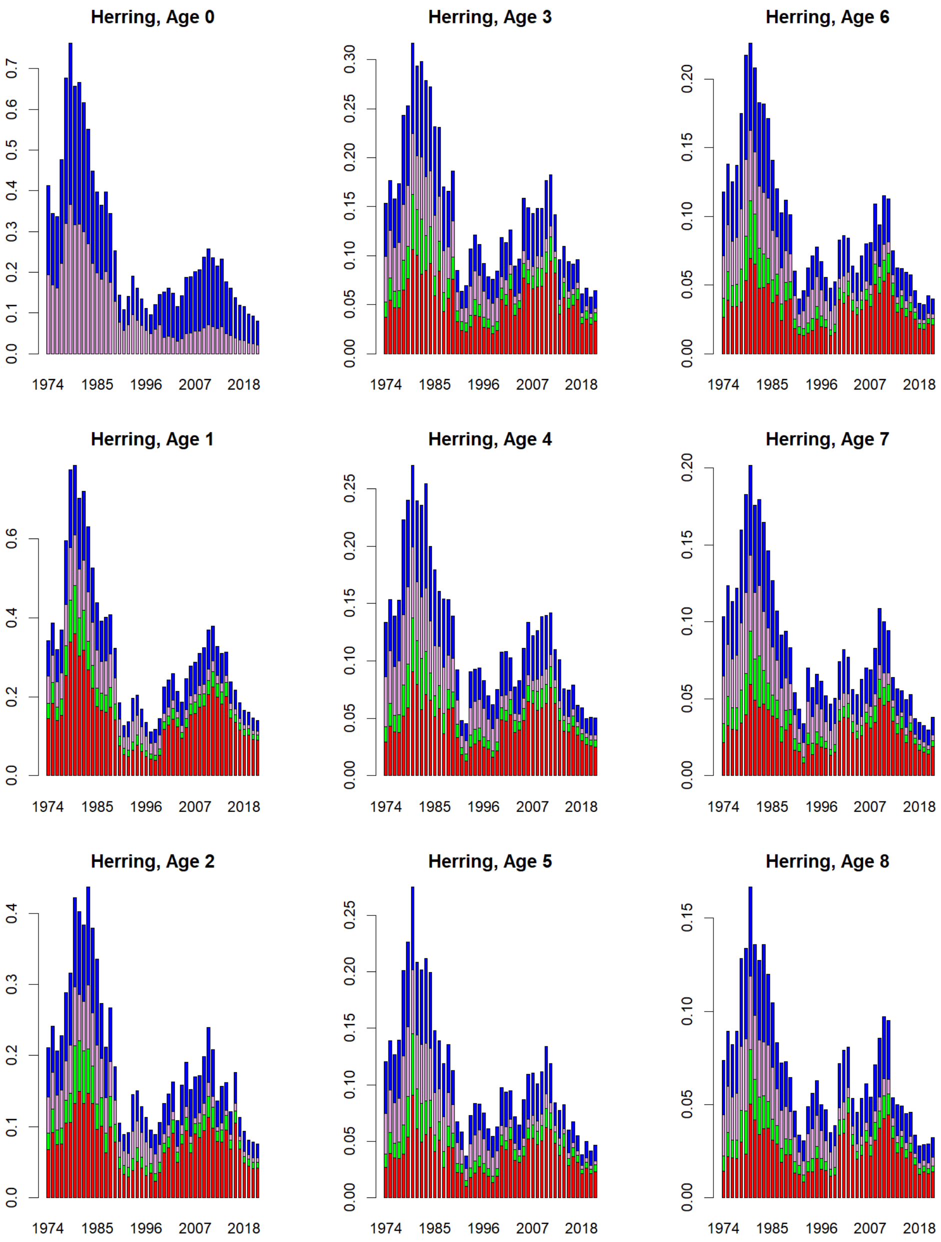


Figure 5.1-9. Annual predation mortality (M2) of herring the colours show M2 by quarter (green Q3, blue Q4, black Q1 and red Q4).

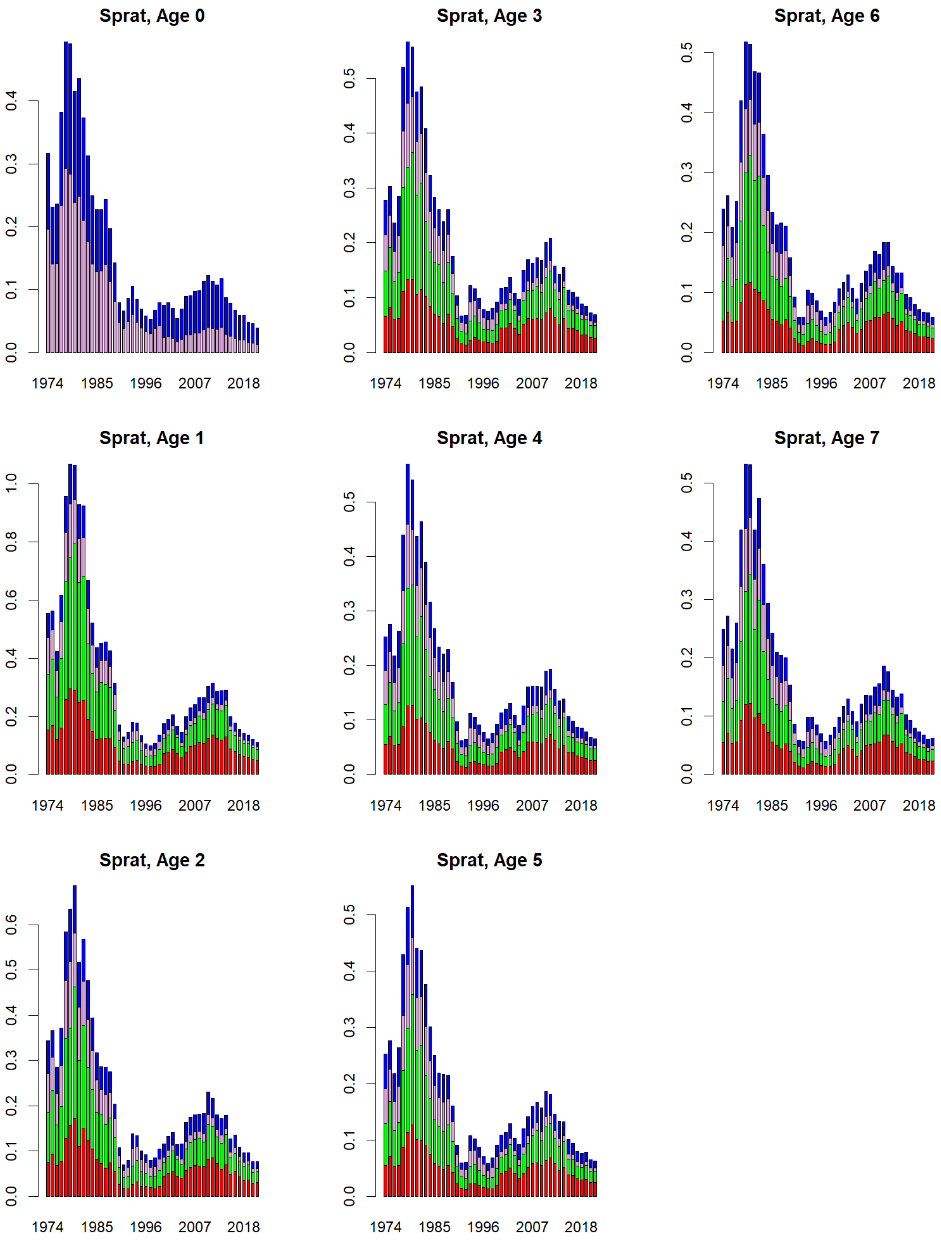


Figure 5.1-10. Annual predation mortality (M2) of herring the colours show M2 by quarter (green Q3, blue Q4, black Q1 and red Q4).

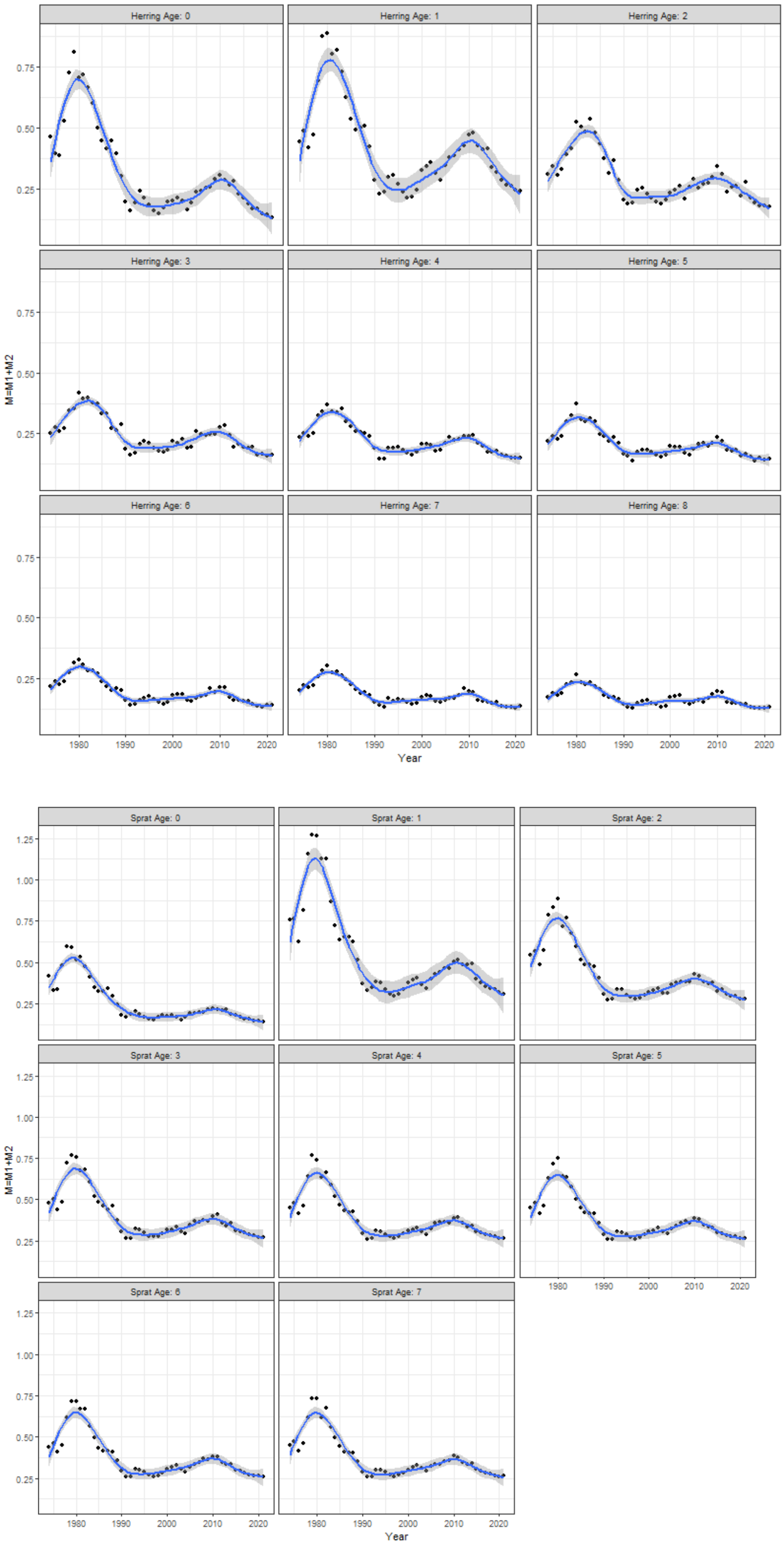


**Figure 5.1-11. Annual natural mortalities (M=M1+M2) by species and age. Black dots are the sum of quarterly M1 and M2; the blue line is a gam spline estimate.**

**Figure 5.1-12. Herring. Comparison of predation mortality (M2) estimated by the 2019 key-run and by the 2022 keyrun.**

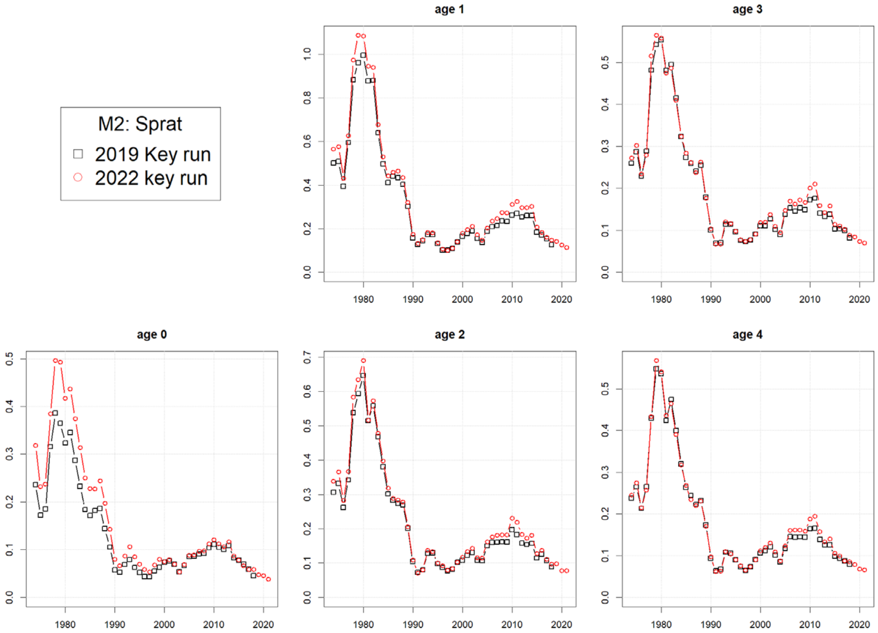


**Figure 5.1-13. Sprat. Comparison of predation mortality (M2) estimated by the 2019 key-run and by the 2022 keyrun.**

### Uncertainties of parameters and output

SMS estimate the uncertainties of selected output variables using the Hessian matrix and the delta-method approximation. Most variables like stock number and F for dynamic species are estimated within the model, while other variables like the stock numbers of the “external predators” cod are assumed known without errors. With cod as the only predator, this combination of estimated and assumed “known” variables will certainly lead to an underestimate of the uncertainties of e.g. predation mortality.

Therefore, the uncertainties estimated from the Hessian matrix are not presented in details.

An example of estimated uncertainties is presented in Figure 5.1-14. The confidence interval seems too tight!

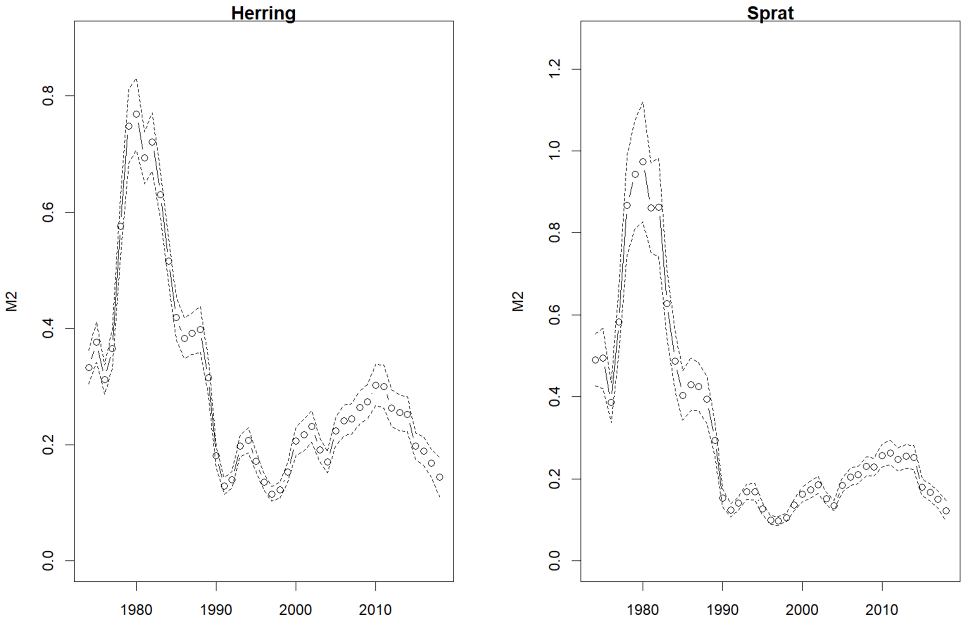


**Figure 5.1-14. Values of M2 and 95% confidence interval (+- 2*standard deviation) for age 1 of herring and sprat. From the 2019 keyrun.**

### Sensitivity test

To get a better idea on true uncertainties several sensitivity runs were carried out as mainly as part of the **2019 and 2022 keyrun**:

1. Retrospective analysis (5 year peel of all input data). 2022 keyrun.
2. Sensitivity to stomach data (old vs. new stomach data set). 2019 keyrun.
3. Sensitivity to stomach data (aggregation stomach data over a 5 or 10 years period). 2019 keyrun.
4. Sensitivity towards using different assumptions for size selection. 2019 keyrun.
5. Sensitivity towards using or not using an overlap index for Other Food. 2019 keyrun.
6. Sensitivity towards consumption rates. 2019 keyrun.
7. Sensitivity towards residual mortality (M1). 2022 keyrun.
8. Comparison of the 2012, 2019 and 2022 keyrun.
9. Comparison with the Gadget model run. 2019 keyrun.

#### Retrospective analysis (5 year peel of all input data)

The retrospective analysis shows variable estimates of recruitment, SSB and F for the terminal years in the time series, (Figure 5.2-1). Comparison with the same kind of output for the ICES assessment (WGBFAS, 2022) reveals however a similar variability in the ICES single species assessment output.

The retrospective analysis show a consistent estimate of predation mortalities (Figure 5.2-2). This consistent estimate is probably also because all runs use the same stomach contents data; the last year with stomach data is 2014. As for all other retrospective assessment analysis, values (M2) in the terminal year of the time-series have larger uncertainties.

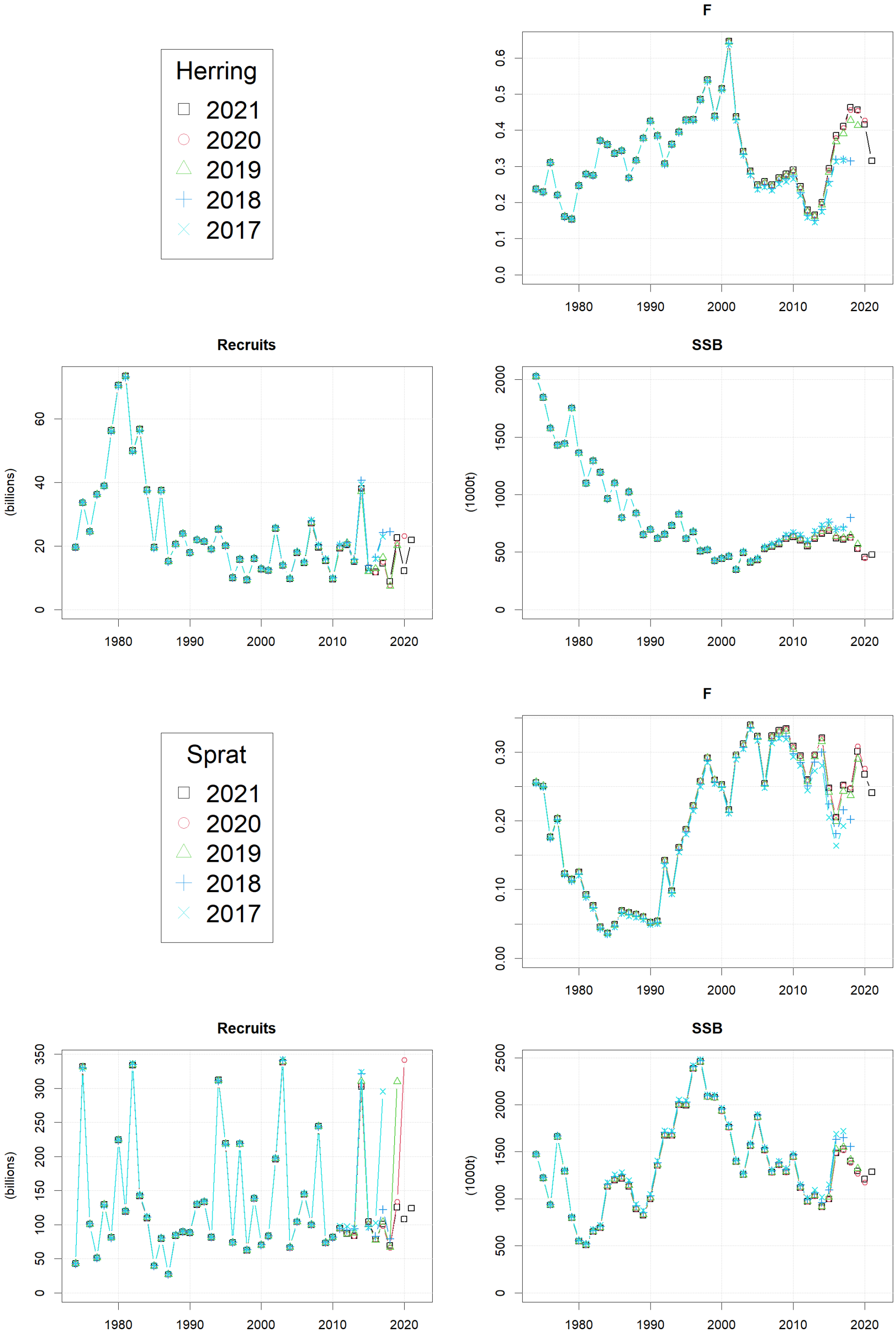


Figure 5.2-1. Retrospective analysis for herring and sprat. Summary output.

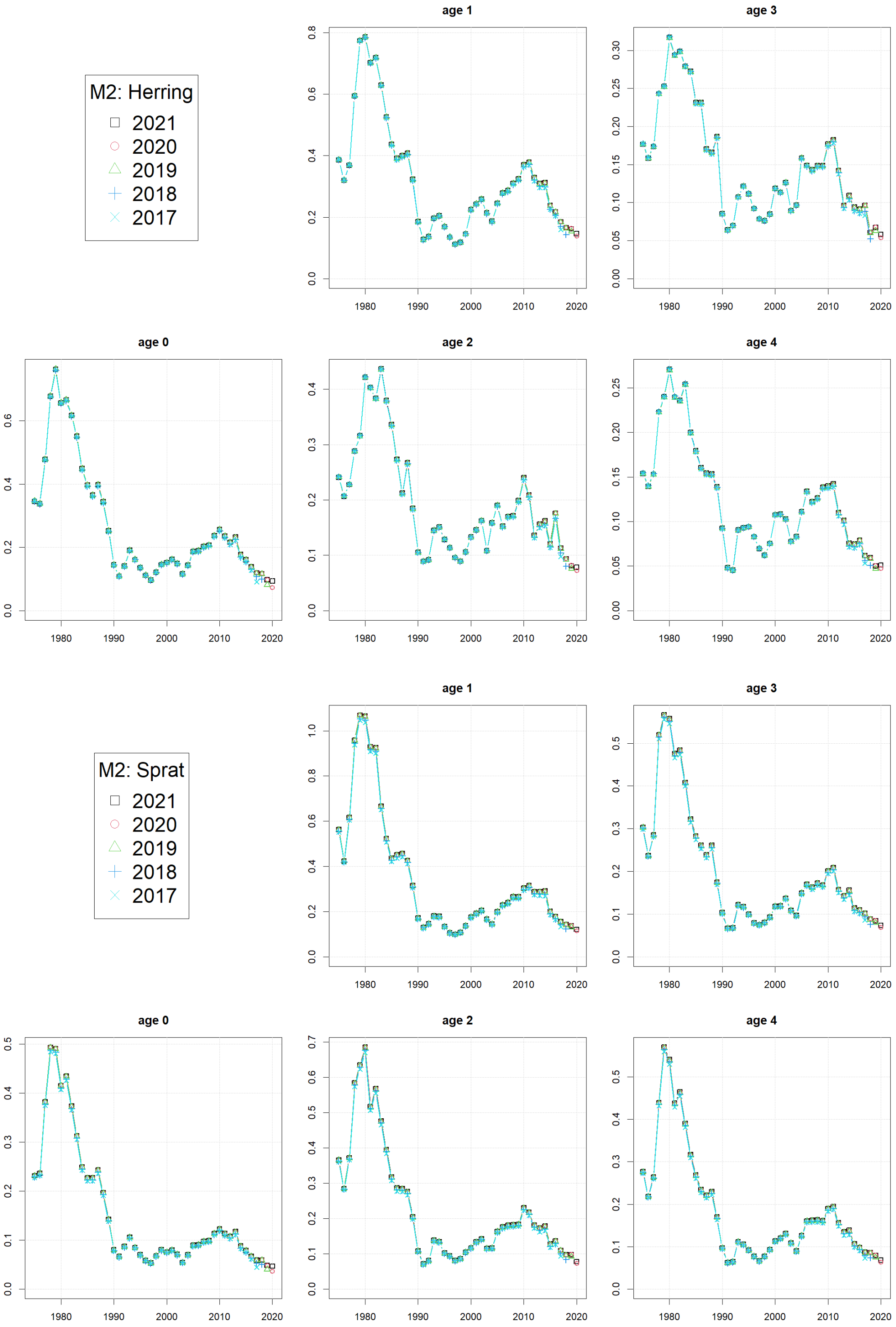


Figure 5.2-2. Retrospective analysis for herring and sprat, M2 at age.

### Sensitivity to stomach data (old vs. new stomach data set)

The choice of stomach contents data, “old”, “new” or combined was investigated as part of the 2019 keyrun. The actual choice has limited effect on the SMS stock summary output (Figure 5.2-4) or predation mortalities (Figure 5.2-4)

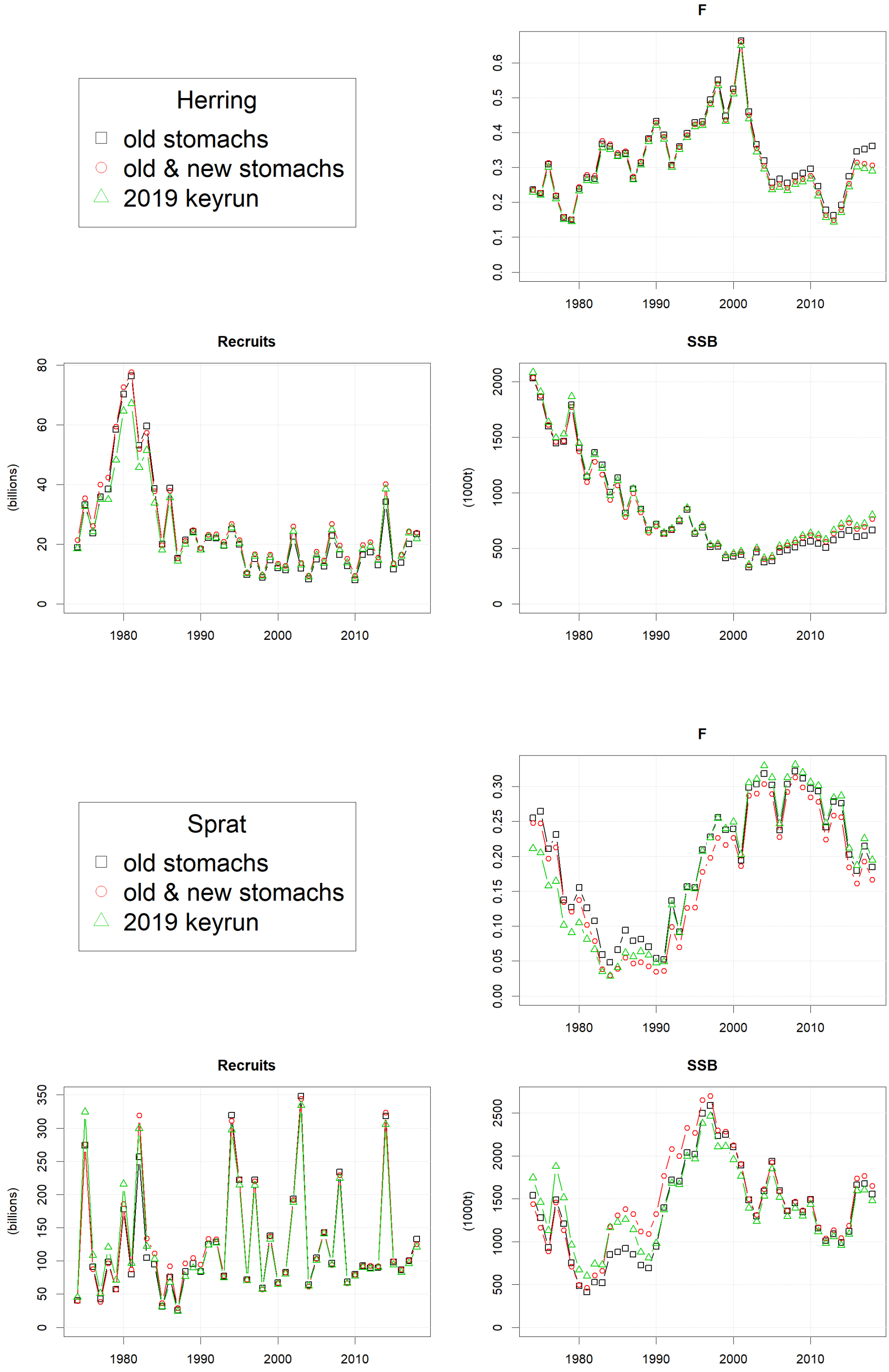


**Figure 5.2-3. Comparison of output from SMS runs with combinations of contents data: “old” pooled stomach data, “new” individually sample stomachs. The keyrun uses only the “new” data.**

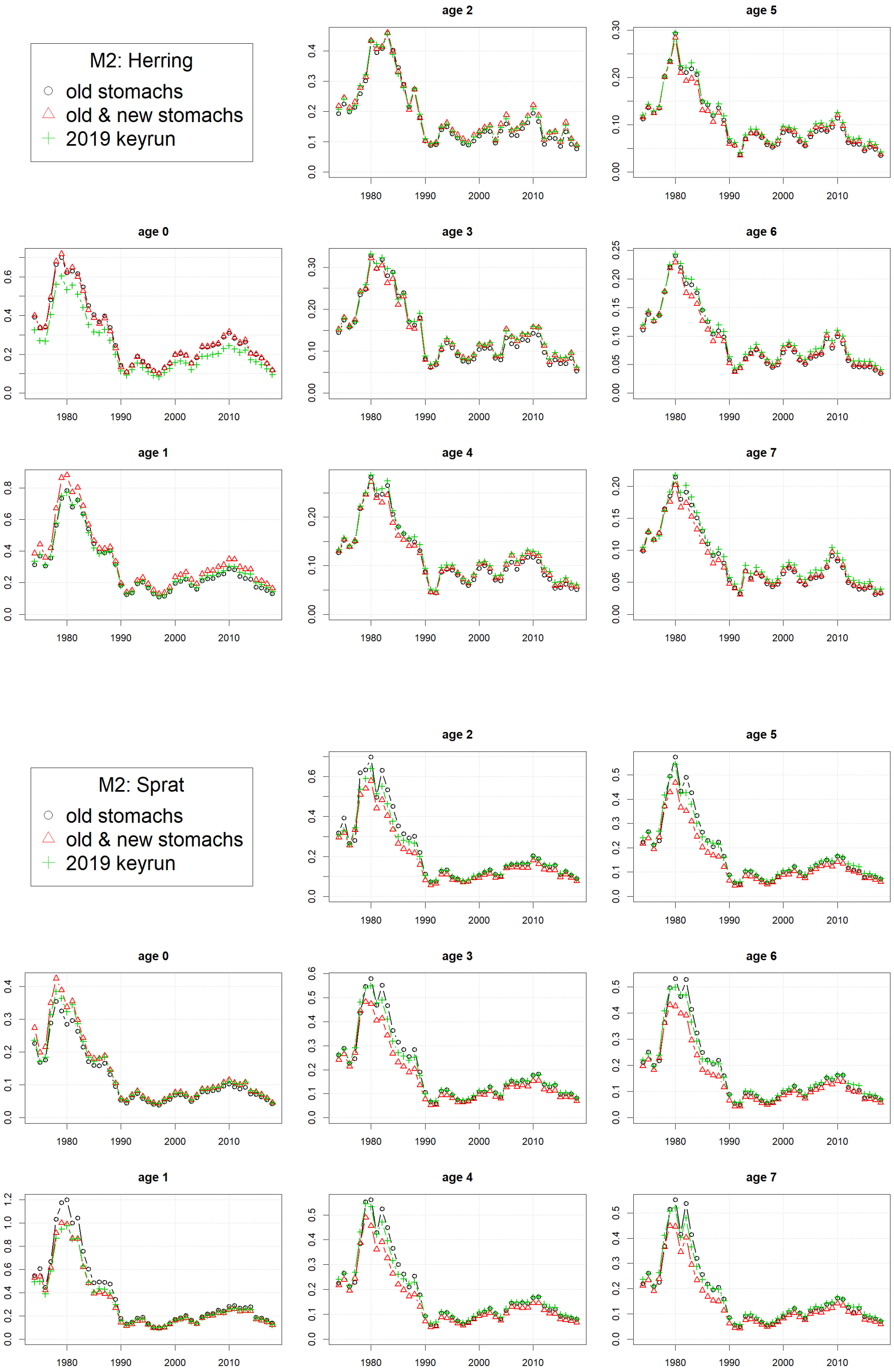


**Figure 5.2-4. Comparison of M2 values from SMS runs with combinations of contents data: "old" pooled stomach data, "new" individually sampled stomachs. The keyrun uses only the "new" data.**

### Sensitivity towards using different assumptions for size selection

Three options for predator prey size selection were tried as part of the 2019 keyrun:

1. Log normal size selection (keyrun): a predator has a preferred prey size ratio and a prey twice as big as the preferred size is as attractive as another half the prey size. The preferred size ration and its variance are estimated by SMS.
2. Uniform size selection: a size preference at 1 within the range of the observed size ratio and 0 outside that ratio.
3. Constraint uniform size selection: as Uniform size selection, but the size preference ratio is constrained to exclude "outliers" from the observed size ratio, estimated from a quantile regression (Figure 5.2-5).

.

The main performance statistics of a SMS run for the three size selection models (Table 5.2-1) show the best model likelihood and AIC for the keyrun.

Stock summary output (Figure 5.2-6) and M2 (Figure 5.2-7) are quite sensitive to the choice of size selection option. It seems as if the "constraint uniform" option excludes interactions from medium sized cod on larger herring (Figure 5.2-5) such that M2 on herring ages 4-8 becomes very low. The (unconstraint) "uniform" options includes the full observed predator/prey size ratio which results in a higher M2 for the older herring than for the "constraint uniform" option.

The "constraint uniform" option performed well in the 2012 keyrun, however there is difference in the quality of stomach contents data used in the old and the new 2019 keyrun. The old keyruns made use of stomach contents data with large size classes for predator preys, e.g. sprat 5-10-15 cm, while the new stomach data uses a much smaller size classes, e.g. by cm group for sprat. With wider size classes, the in predator/prey size ratio becomes imprecise, such that the cutting of "outliers" by the "constraint uniform" options had a limited effect. With the new data, the full range of observations should probably be used, if a uniform size selection option is used.

**Table 5.2-1. SMS main performance statistics from a SMS run with the "uniform size selection", "constraint uniform size selection" and the keyrun.**

## uniform size selection

```
objective function (negative log likelihood):  -1166.15
Number of parameters: 289
Number of observations used in likelihood: 14892
Akaike information criterion (AIC):   -1754.31

Number of observations used in the likelihood:
unweighted objective function contributions (total):
                Catch    CPUE    S/R   Stom.       Sum
Cod               0.0     0.0    0.0  -173.4      -173
Herring        -667.6  -117.3  -12.5     0.0      -797
Sprat          -100.9  -106.0   -5.5     0.0      -212
Sum            -768.6  -223.3  -18.0  -173.4     -1183
```

## constraint uniform size selection

```
objective function (negative log likelihood):  -1110.12
Number of parameters: 289
Number of observations used in likelihood: 14892
Akaike information criterion (AIC):   -1642.24

unweighted objective function contributions (total):
                Catch    CPUE    S/R   Stom.       Sum
Cod               0.0     0.0    0.0  -182.5      -183
Herring        -645.5  -121.1   -8.2     0.0      -775
Sprat           -56.8  -103.5   -5.2     0.0      -166
Sum            -702.4  -224.7  -13.4  -182.5     -1123
```

## Log-normal size selection (keyrun 2019)

```
objective function (negative log likelihood):  -1232.3
Number of parameters: 292
Number of observations used in likelihood: 14892
Akaike information criterion (AIC):   -1880.6

unweighted objective function contributions (total):
                Catch    CPUE    S/R   Stom.       Sum
Cod               0.0     0.0    0.0  -256.2      -256
Herring        -660.4  -118.7   -8.6     0.0      -788
Sprat           -92.3  -104.0   -5.6     0.0      -202
Sum            -752.7  -222.7  -14.2  -256.2     -1246
```

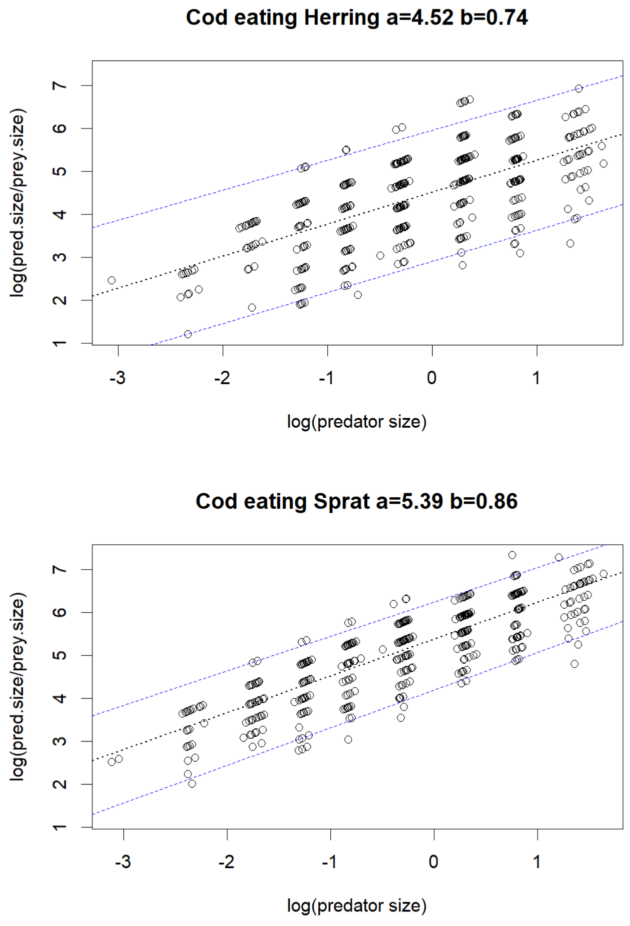


**Figure 5.2-5. Quantile regression with observations of predator and predator/prey sizes. The blue lines shows the 2.5% and 97.5 % percentile lines, which defines the "size selection window".**

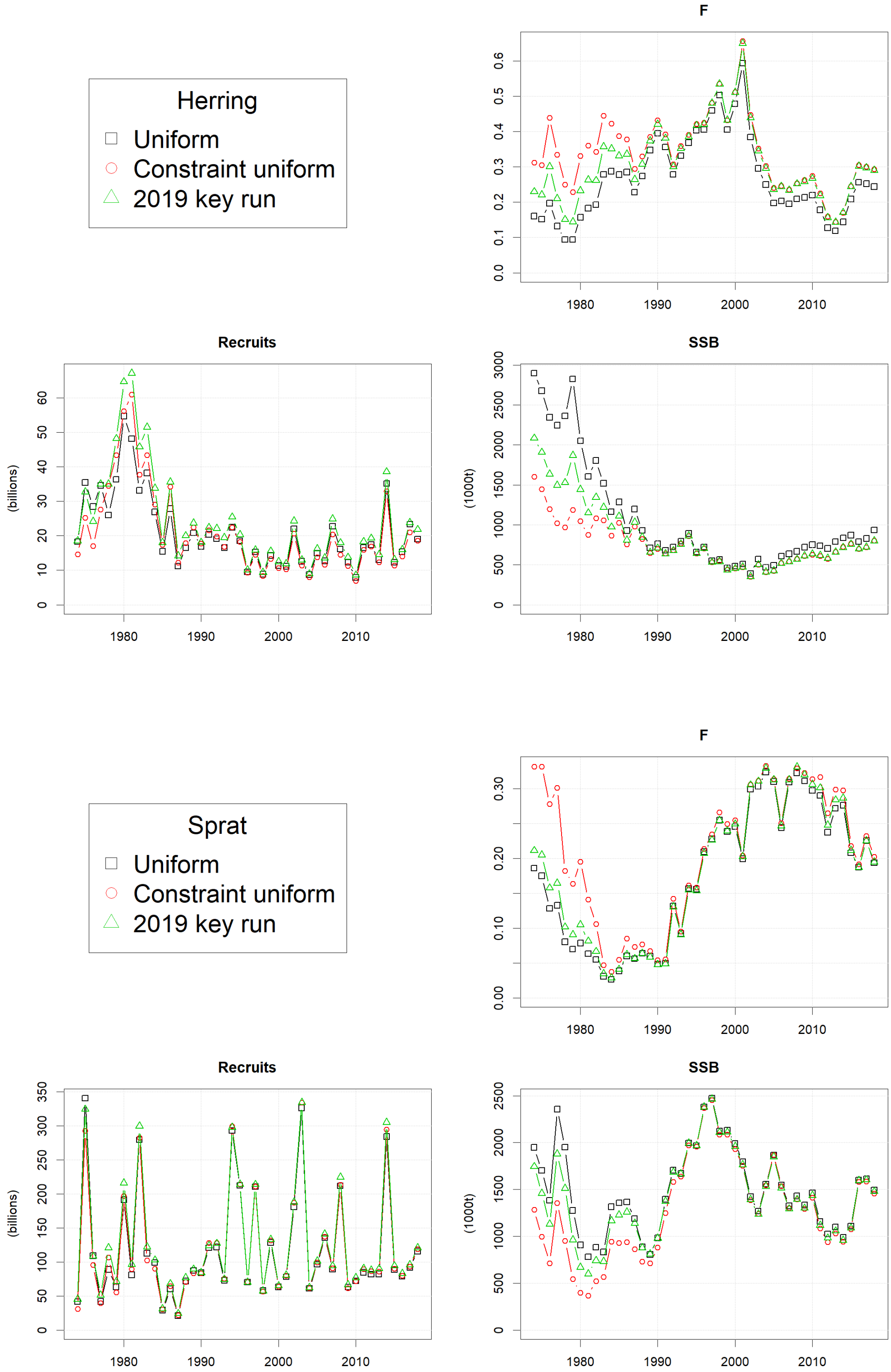


**Figure 5.2-6. Comparison of output from SMS runs with three options for predator prey size selection.**

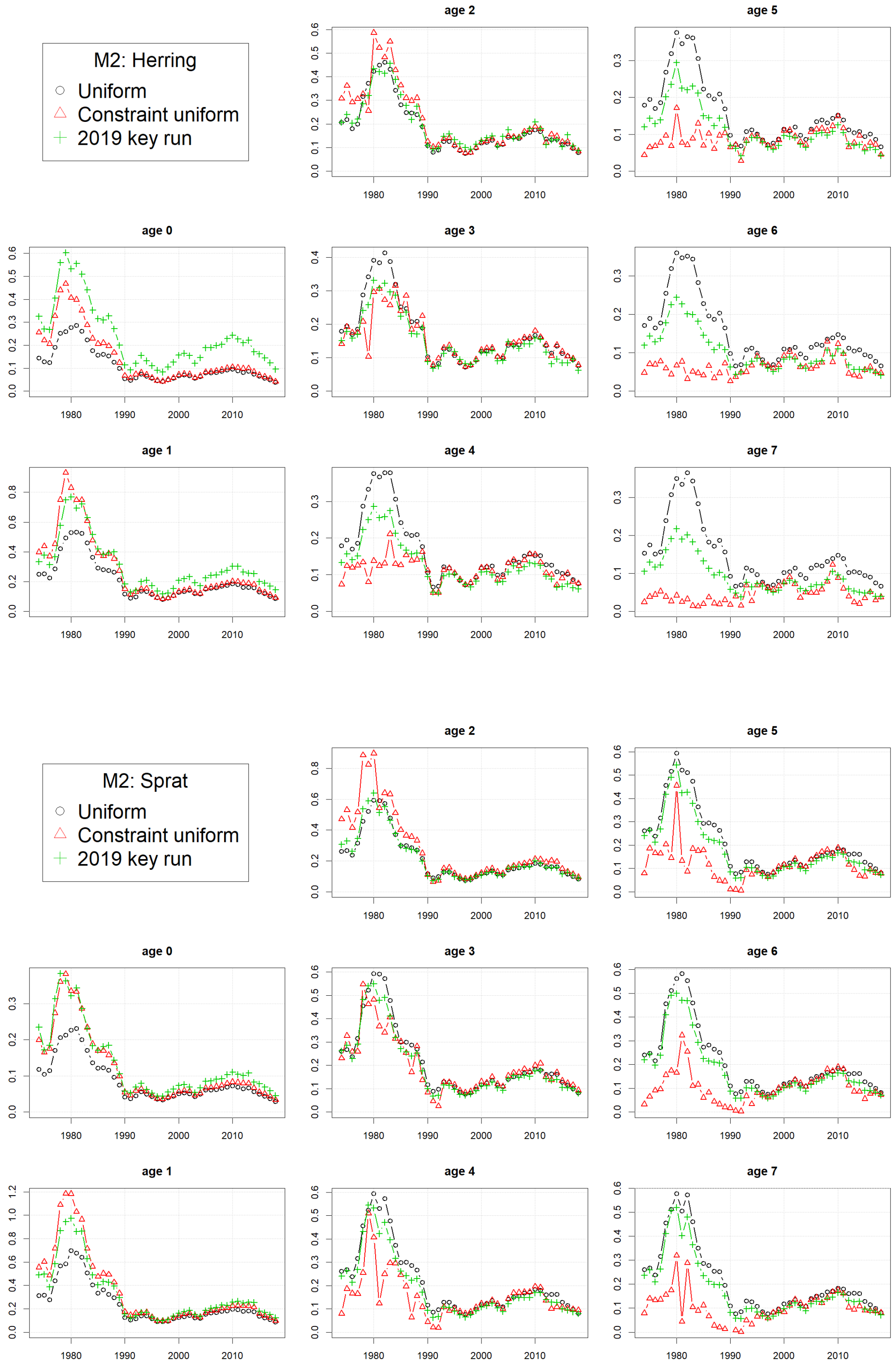


**Figure 5.2-7. Comparison of M2 values from SMS runs with three options for predator prey size selection.**

### Sensitivity to stomach data (aggregation over a 5 or 10 years period)

Stomachs data are by default aggregated for each combination of year and quarter of the year. For most cases this leads to a rather few stomachs for some of the predator length classes and uncertainties on how partly identified prey items should be assigned. Aggregating stomach data over more years, e.g. 5 or 10 years, provides a larger sample size and a smaller observation uncertainties, but an e.g. ″10 years diet″ will not reflect the variability in available food for the individual years. This effect of aggregating stomachs over several years was investigated as part of the 2019 keyrun.

The average likelihood contribution (Table 5.2-2) show that likelihood per stomach contents observation becomes better (more negative) when data are aggregated over some years compared to the keyrun which uses data by year. The best average likelihood for stomach data is obtained using a 5-years aggregation. This may be interpreted that pooling stomach data between year gives a higher precision (more stomachs) of data used by SMS. However, using a very wide year range may negatively affect the fit between "observed" stomach contents and the model estimate of stomach contents calculated for the midpoint of the years used in the data aggregation. Likelihood contributions from For Catch, CPUE and S/R observations are quite the same for the three configurations.

M2 values for the three configurations are differ mainly for age 0 and 1 of herring and sprat (Figure 5.2-8).

**Table 5.2-2. Objective function contributions (per observation) from SMS models using stomach contents data aggregated over 5, 10 years and from the keyrun.**

**5 years aggregation:**

| | Catch | CPUE | S/R | Stomachs |
|---|---|---|---|---|
| Cod | 0.00 | 0.00 | 0.00 | -0.28 |
| Herring | -0.46 | -0.46 | -0.18 | 0.00 |
| Sprat | -0.08 | -0.33 | -0.12 | 0.00 |

**10 years aggregation:**

| | Catch | CPUE | S/R | Stomachs |
|---|---|---|---|---|
| Cod | 0.00 | 0.00 | 0.00 | -0.26 |
| Herring | -0.47 | -0.45 | -0.18 | 0.00 |
| Sprat | -0.09 | -0.33 | -0.11 | 0.00 |

**Keyrun:**

| | Catch | CPUE | S/R | Stomachs |
|---|---|---|---|---|
| Cod | 0.00 | 0.00 | 0.00 | -0.16 |
| Herring | -0.46 | -0.47 | -0.19 | 0.00 |
| Sprat | -0.07 | -0.33 | -0.12 | 0.00 |

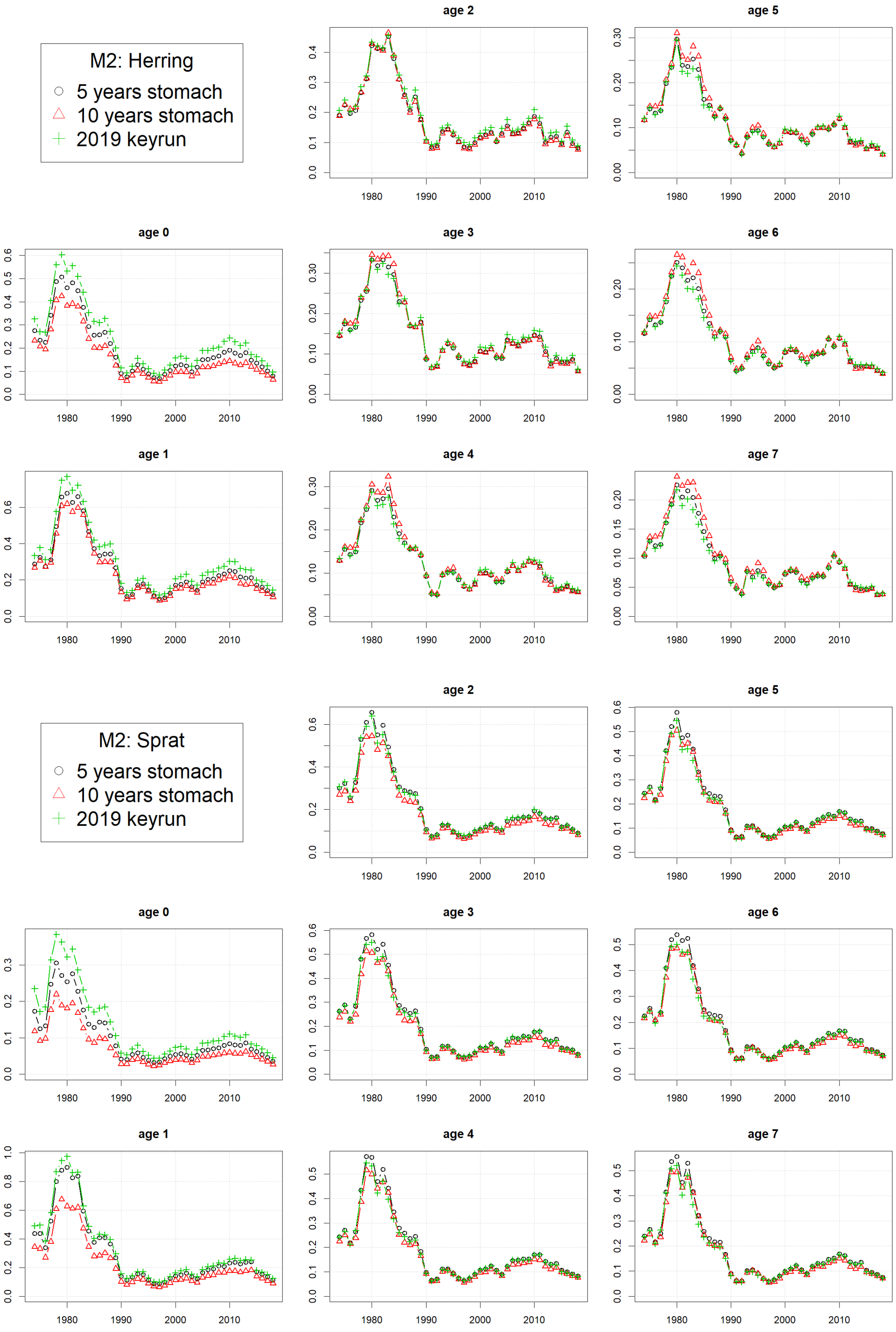


**Figure 5.2-8. M2 estimated from SMS runs using stomach contents data aggregated over 5, 10 years and from the keyrun.**

#### Sensitivity towards using an overlap index for Other Food

The “other food” prey *Saduria entomon* is an important benthic prey item for cod. The occurrence of this prey depends on the oxygen level at the bottom, which have changed considerably in the model timespan.

A time-series of total area ($km^2$) of hypoxic bottoms (between 20 and 100 m depth) was used to develop an index for overlap with *Saduria entomon* and other benthic components, assuming ≤1 ml $l^{-1}$ (approx. 1.4 mg $l^{-1}$) as threshold for oxygen concentration to indicate failure in benthic productivity. With $A_h$ indicating the hypoxic bottom area, the index was defined as $(A_h / maxA_h)^{-1}$, yielding higher values the smaller the hypoxic area was in a given year. We applied a 5-yr running mean. Weighting the areas with the sub-division specific cod distribution did not change the index time series except for the last two years with data, 2013 and 2014.

The overlap index between cod and the prey species herring and sprat was left unchanged (assumed 1 throughout the period)

The performance statistics (Table 5.2-3) for the runs with input overlap index and the 2019 keyrun are almost the same, even though the keyrun has a better total model likelihood. The likelihood contributions from stomach observations are the same for the two models.

Stomach contents residuals (Figure 5.2-9) are similar to the keyrun residuals (Figure 5.1-4) but residuals are actually less clustered in positive and negative residuals when the input overlap index is applied.

**Table 5.2-3. SMS main performance statistics from a SMS run with input overlap index for Other Food and the keyrun.**

## Log-normal size selection (2019 keyrun)

```
objective function (negative log likelihood):  -1232.3
Number of parameters: 292
Number of observations used in likelihood: 14892
Akaike information criterion (AIC):   -1880.6
unweighted objective function contributions (total):
                Catch     CPUE     S/R   Stom.       Sum
Cod               0.0      0.0     0.0  -256.2      -256
Herring        -660.4   -118.7    -8.6     0.0      -788
Sprat           -92.3   -104.0    -5.6     0.0      -202
Sum            -752.7   -222.7   -14.2  -256.2     -1246
```

## With input overlap index for Other Food

```
Objective function (negative log likelihood):  -1210.41
Number of parameters: 292
Number of observations used in likelihood: 14892
Akaike information criterion (AIC):   -1836.83

unweighted objective function contributions (total):
                Catch     CPUE     S/R   Stom.       Sum
Cod               0.0      0.0     0.0  -256.3      -256
Herring        -662.2   -116.4   -20.4     0.0      -799
Sprat           -71.9   -102.3    -5.9     0.0      -180
Sum            -734.1   -218.7   -26.3  -256.3     -1235
```

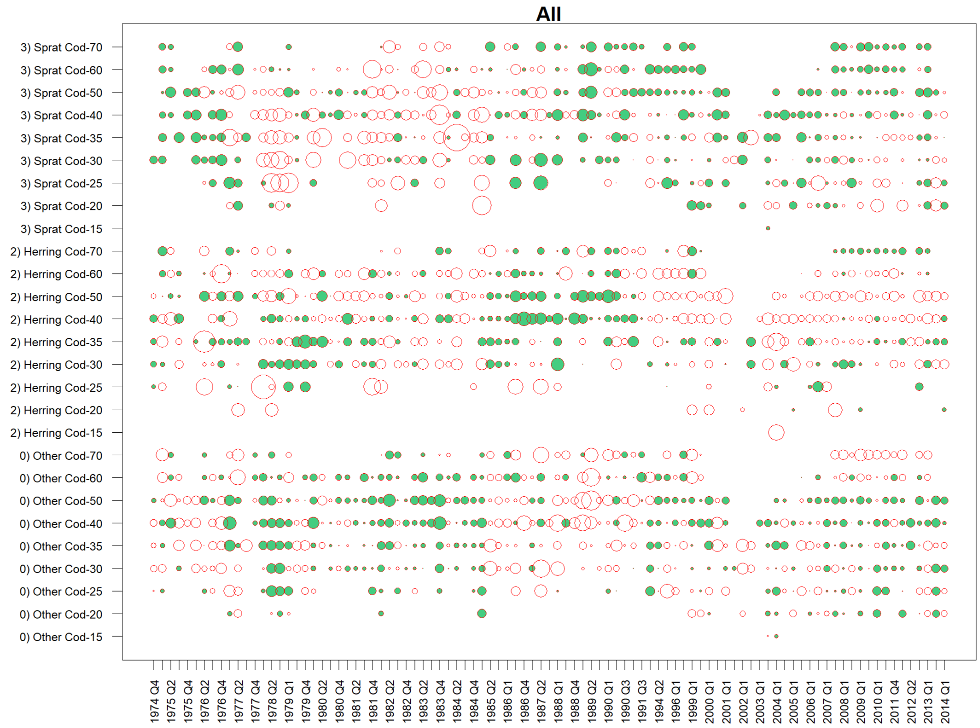


**Figure 5.2-9. Stomach contents residuals ("Dirichlet residuals", Peter Lewy, pers. comm.). The y-axis show prey group and predator (cod) size class. The x-axis is time period sorted by year and quarter. Green dots show that the observed stomach contents are lower than the model estimate.**

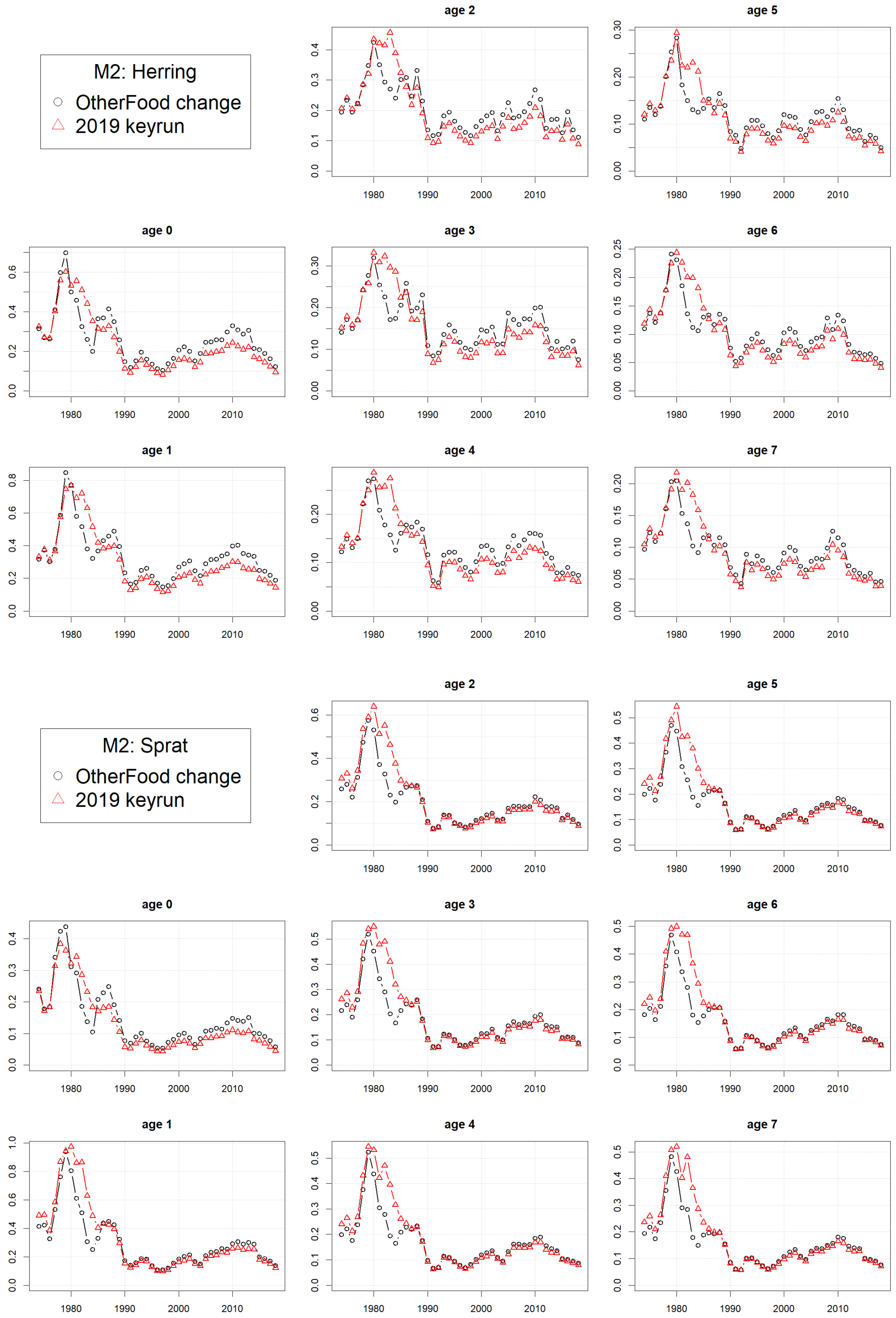


**Figure 5.2-10. M2 estimated from a SMS run with input overlap index for Other Food and from the keyrun.**

### Sensitivity towards consumption rates

SMS can estimate a scaling factor for the input consumption rate by species. For cod this was estimated to 0.47, with a standard deviation at 0.07 for the 2019 keyrun. This run was used to illustrate the sensitivity of M2 to consumption rates.

The performance statistics (Table 5.2-4 and Table 2.4-1) show a slightly better fit, when the factor to the input consumption rate is applied. Likelihood contribution from stomach becomes better on the cost of the likelihood for catch at age. A similar exercise using the 2022 keyrun estimated a ration scaling factor at 0.46 (sd 0.07) and the same shifts in likelihood contributions as shown for the 2019 keyrun.

M2 values are lower when a considerably lower consumption rate are applied, but the reduction is not linear to the reduction in consumption, as expected (Figure 5.2-11). The reduction in M2 is larger for herring than for sprat.

**Table 5.2-4 SMS main performance statistics from a SMS run with input overlap index for Other Food and the 2019 keyrun.**

## Log-normal size selection (2019 keyrun)

```
objective function (negative log likelihood):  -1232.3
Number of parameters: 292
Number of observations used in likelihood: 14892
Akaike information criterion (AIC):   -1880.6
unweighted objective function contributions (total):
                Catch    CPUE     S/R   Stom.       Sum
Cod              0.0     0.0     0.0  -256.2      -256
Herring       -660.4  -118.7    -8.6     0.0      -788
Sprat          -92.3  -104.0    -5.6     0.0      -202
Sum           -752.7  -222.7   -14.2  -256.2     -1246
```

## With input consumption rates *0.47

```
objective function (negative log likelihood):  -1244.78
Number of parameters: 293
Number of observations used in likelihood: 14892
Akaike information criterion (AIC):   -1903.57

unweighted objective function contributions (total):
                Catch    CPUE     S/R   Stom.       Sum
Cod              0.0     0.0     0.0  -273.5      -273
Herring       -659.6  -119.9   -16.1     0.0      -796
Sprat          -87.2  -103.6    -5.4     0.0      -196
Sum           -746.7  -223.5   -21.5  -273.5     -1265
```

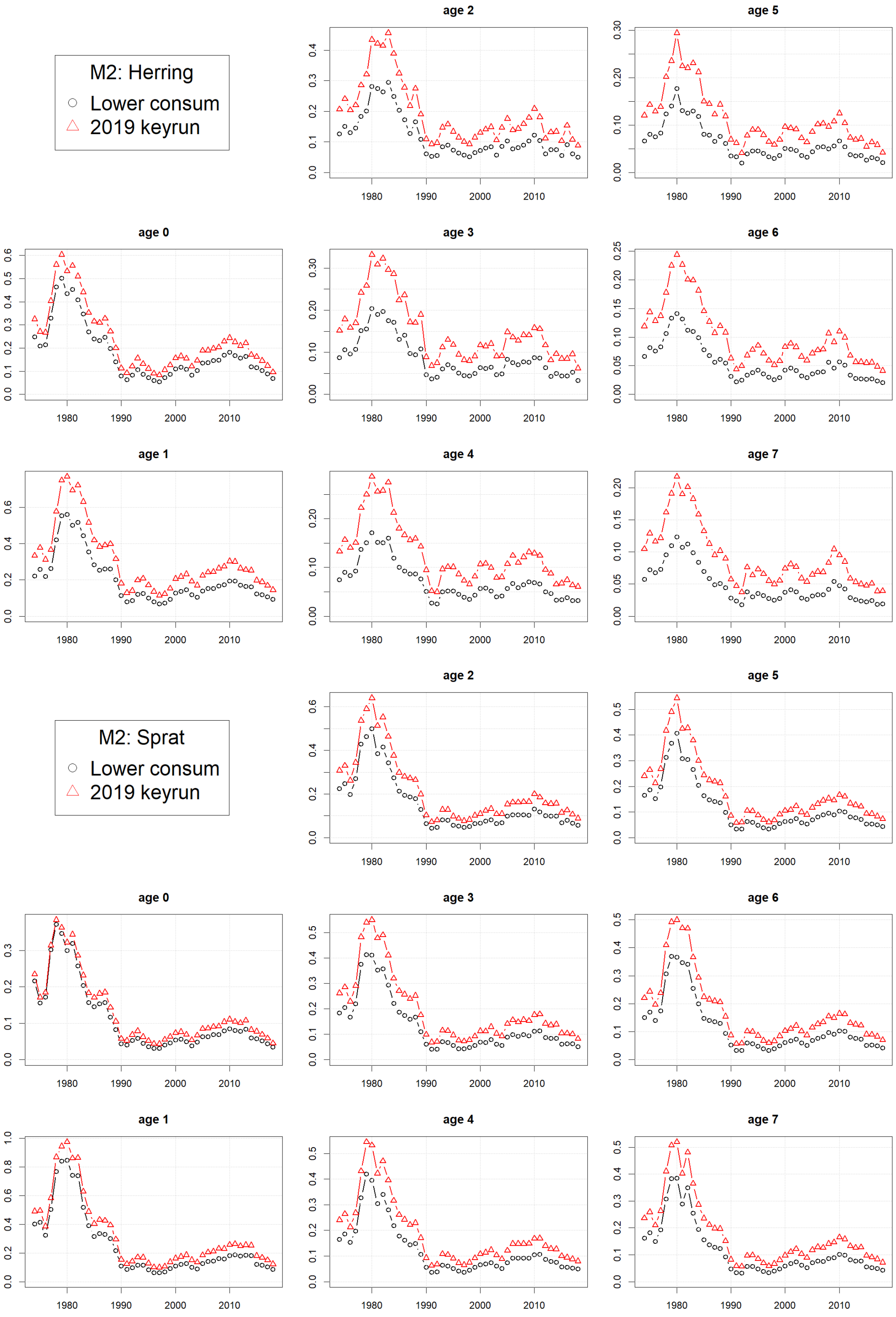


Figure 5.2-11. M2 estimated from a SMS run with lower (47%) consumption rates the keyrun.

**Sensitivity towards residual natural mortality (M1).**

M1 values are assumed to be constant, independent of age and known without errors, however there are no data to actually support the exact values chosen. As part of the 2019 keyrun it was decided to use 0.1 for herring and 0.2 (annual values) for sprat, with the justification that this are the values used for the same species for the North Sea keyrun.

To show the sensitivity of the M1 values on the estimated M2 values, an SMS run was made where the M1 used in the 2022 keyrun were multiplied 0.5.

**Table 5.2-5. SMS main performance statistics from a run with M1 values are multiplied by 0.50**

## M1 halved

```
M1 for herring changed from 0.1 to 0.05 (annual values)
M1 for herring changed from 0.2 to 0.1 (annual values)

objective function (negative log likelihood):  -1614.62
Akaike information criterion (AIC):   -2571.23

unweighted objective function contributions:
                 Catch       CPUE      SSB/R   stomach       Sum
Cod                0.0        0.0        0.0    -320.4      -320
Herring         -855.0     -141.6       -3.9       0.0     -1001
Sprat           -179.9     -117.0       -7.7       0.0      -305
Sum            -1035.0     -258.7      -11.5    -320.4     -1626
```

## Keyrun 2022

```
objective function (negative log likelihood):  -1646.5
Akaike information criterion (AIC):   -2635

unweighted objective function contributions:
                 Catch       CPUE      SSB/R   stomach       Sum
Cod                0.0        0.0        0.0    -337.6      -338
Herring         -857.6     -144.3       -6.5       0.0     -1008
Sprat           -187.9     -118.5       -7.6       0.0      -314
Sum            -1045.4     -262.7      -14.1    -337.6     -1660
```

The model statistics are not affected much by the choice of M1, but the fits for both catch, cpue and stomach data are the best for the keyrun (unchanged M1 values). The model configuration is the same for the two runs, and is chosen on the basis of the keyrun, which may influence this result.

A lower M1 value results in a lower recruitment, a higher F and a lower SSB for both herring and sprat (Figure 5.2-12). The larges changes are seen for sprat with the largest absolute change in M1 (from 0.2 to 0.1). M2 values (Figure 5.2-13) are estimated higher with a lower M1 as the stock sizes of the prey stock are lower and cod has the same total consumption irrespective of M1. Cod, is an "other predator" with input stock size is not affected by the change in M1. The sum of M1 and M2 (M), as used by the single species assessment is practically the same for herring. For sprat, with an higher absolute change in M1, a slightly lower M are estimated for a lower M2 and a slightly higher M2.

To conclude: Estimated values of recruitment, F and SSB depends on the value of M1. The estimated M2 depends on the applied M1 values as well, with the highest M2 for the lowest M1. The sum of M1 and estimated M2 (M) as used in the ICES single species assessment becomes in most cases lower with a lower M1. As the stock size of the predator cod in this run is not affected by the change in M1, the results from the Baltic Sea run may not be valid for e.g. the North Sea, where a change in M1 will also affect the main part of the predators.

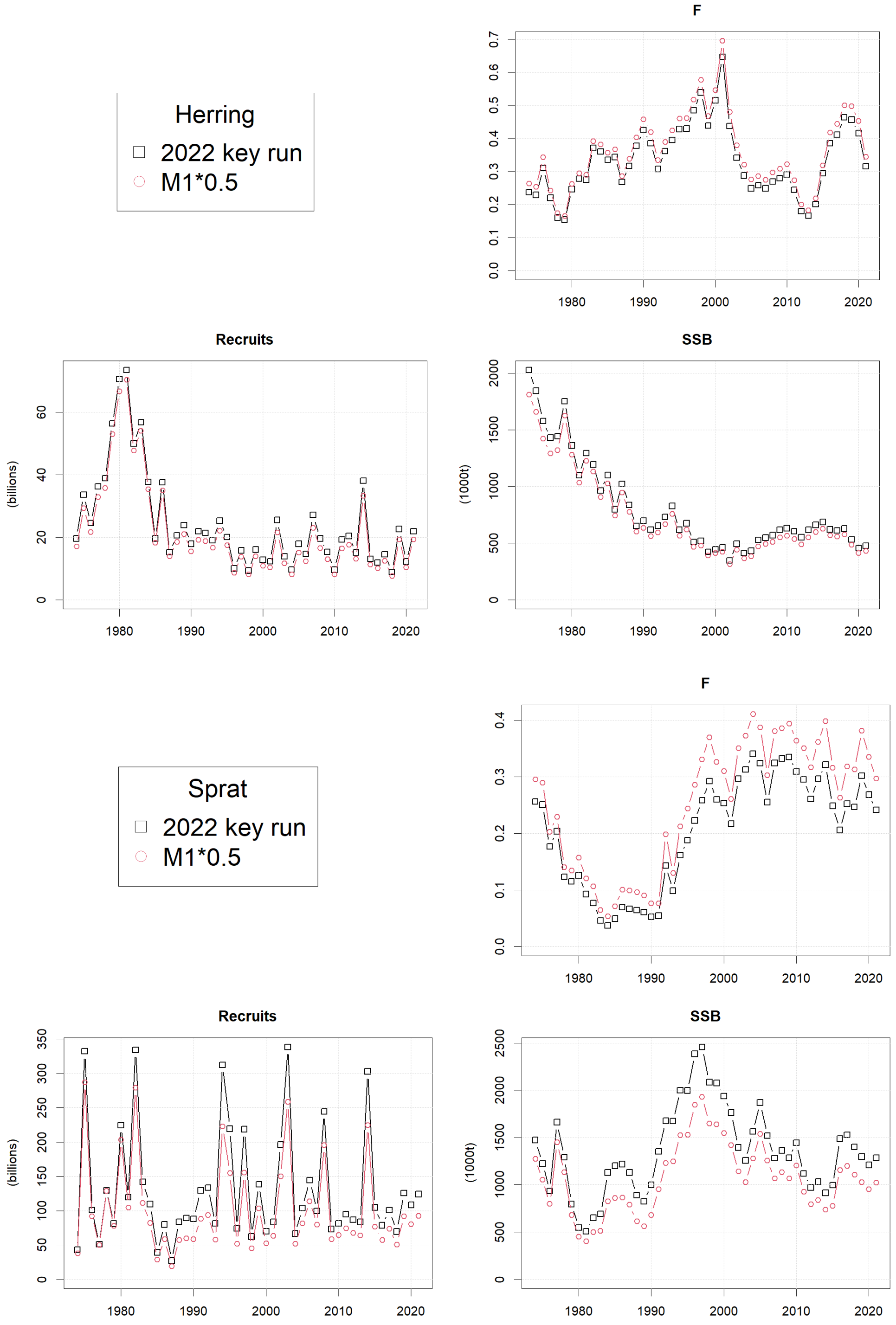


**Figure 5.2-12 Comparison of the 2022 keyrun and a run with M1 multiplied by 0.5. Summary output.**

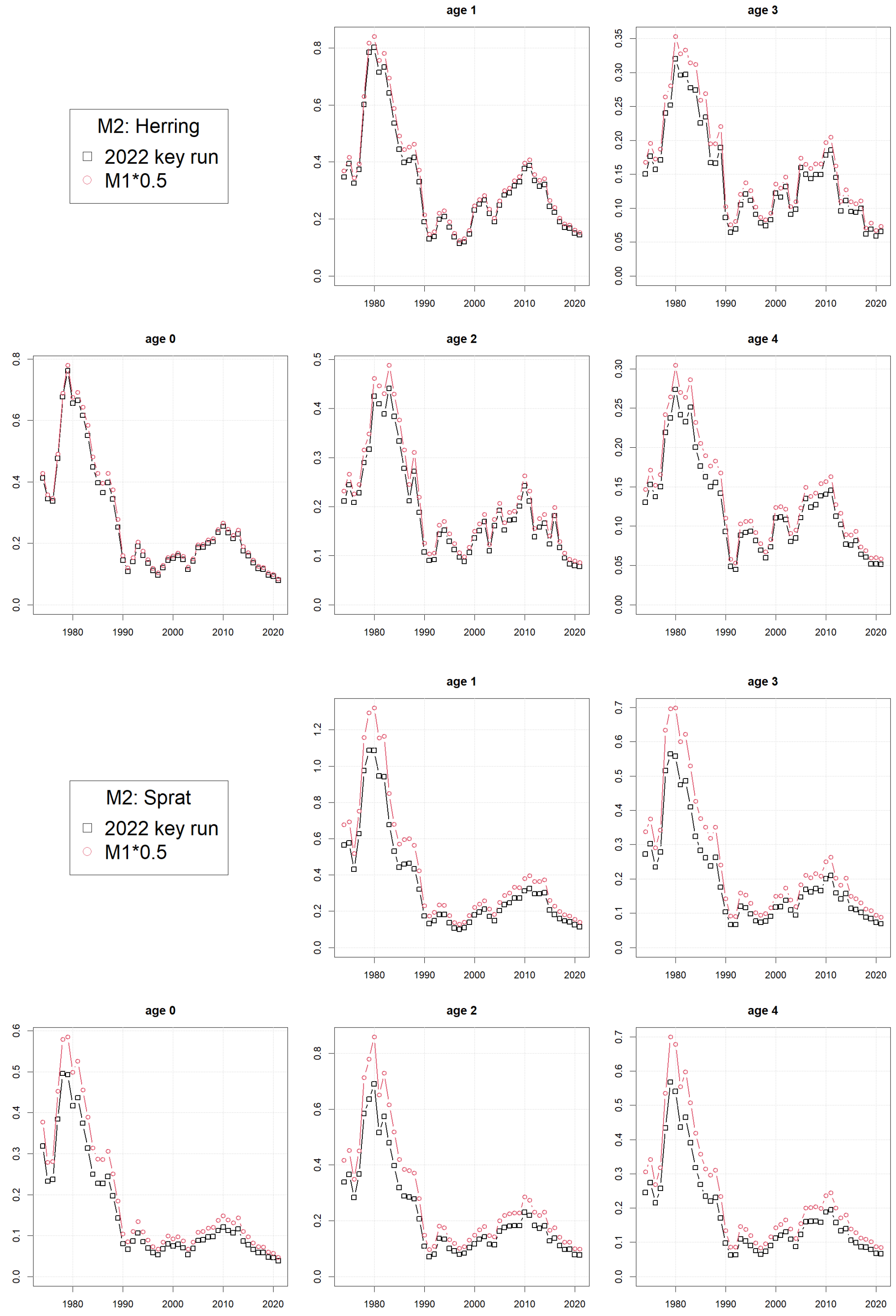


Figure 5.2-13 Comparison of the 2022 keyrun and a run with M1 multiplied by 0.5. Predation mortality (M2).

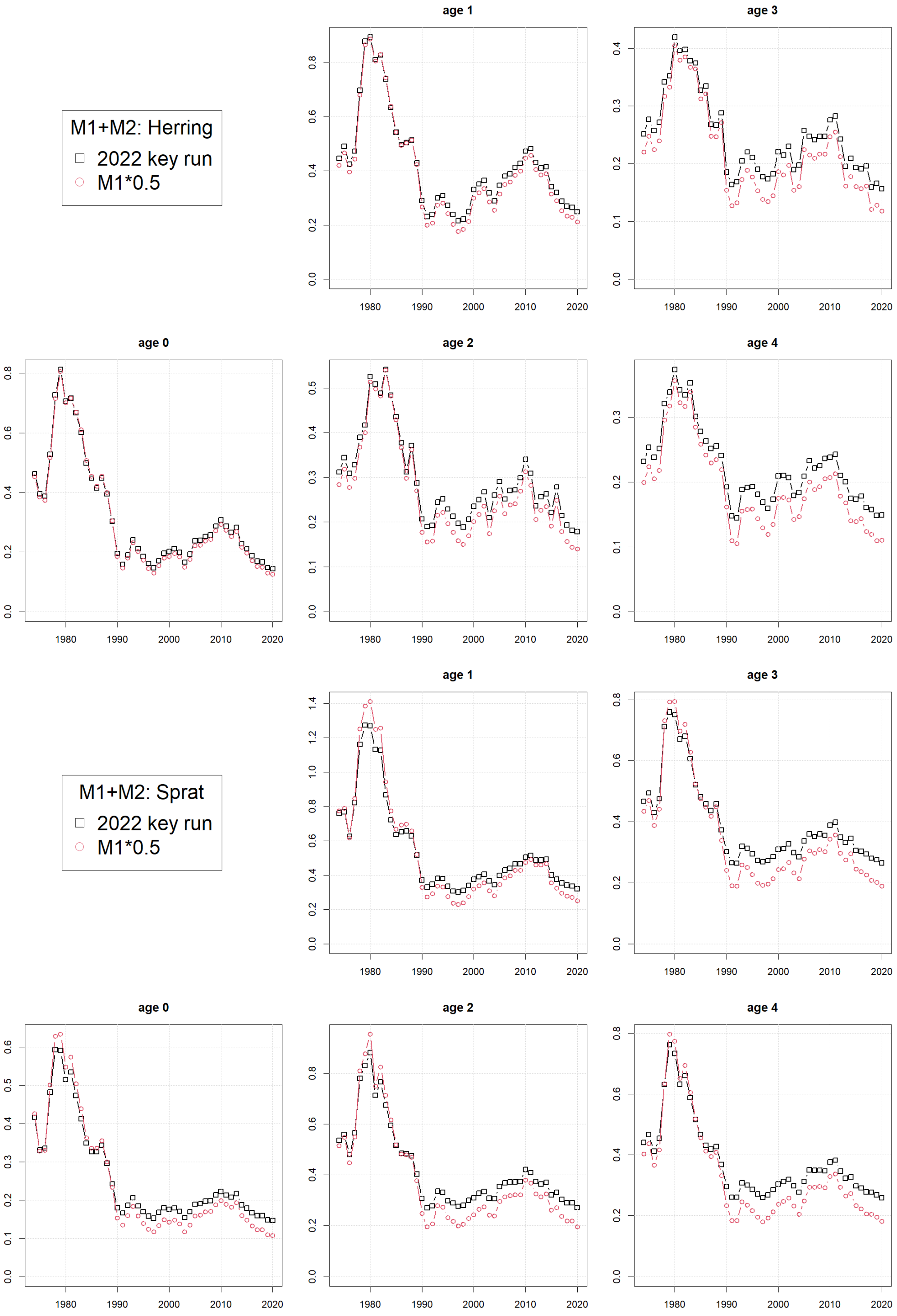


**Figure 5.2-14 Comparison of the 2022 keyrun and a run with M1 multiplied by 0.5. Natural Mortality (M = M1 + M2).**

## Comparison with the old keyruns

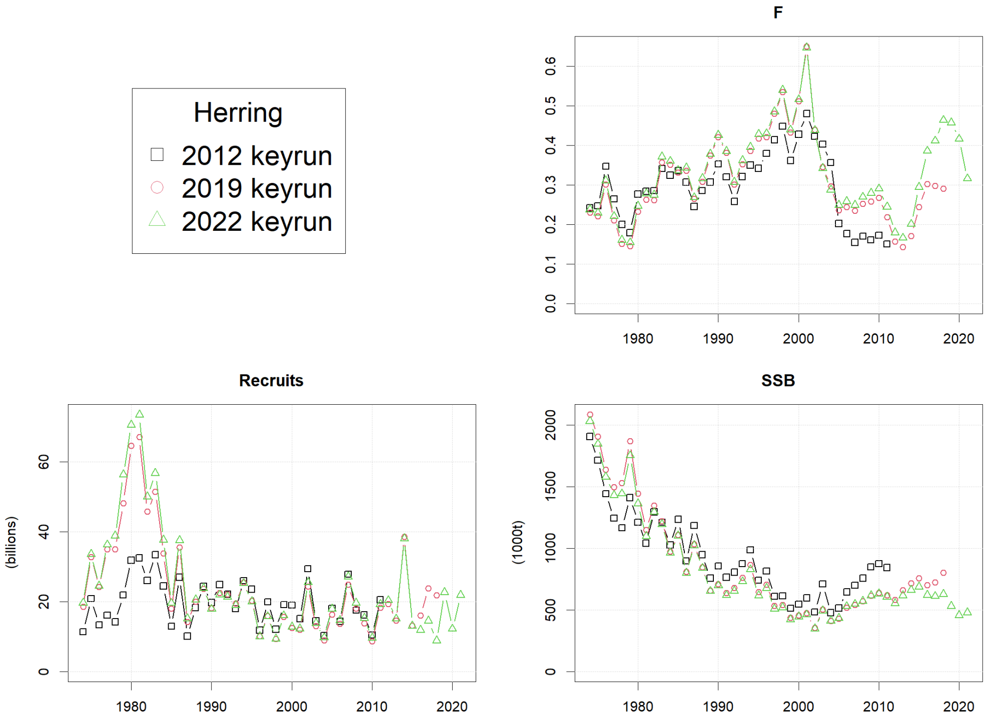
### Comparison of the 2012 and 2019 keyruns

Even though the 2012 and the 2019 keyrun are based on different stomach data, different assumption about the only predator species and different M1, the two keyruns shows quite similar results for the summary output recruitment, SSB and mean F (Figure 5.2-15). Herring F and SSB are similar, while recruitment is considerably higher in the beginning of the time series in the 2019 keyrun, probably as an effect of the assumed larger cod stock. For Sprat, the trend in SSB and F is the same in the two runs, but F in the 2019 keyrun is consistently estimated lower and SSB higher.

The difference in M2 for the two runs is more pronounced, especially for herring (Figure 5.2-16). Herring M2 is now estimated higher for all ages, and much higher for the first part of the time series. The difference is probably due to the assumption of a larger cod stock (especially of larger cod) in the 2019 keyrun, and the application of the predator-prey size selection model in the new keyrun, whereas the old version used a “constraint uniform” size selection. Herring M2 follows better the stock size of cod in the new run which may indicate that the uniform size selection option was not the best choice for the 2012 keyrun.

### Comparison of the 2019 and 2022 keyruns.

The results from the 2019 and 2022 are very similar (Figure 5.2-15 and Figure 5.2-16), as few changes were made in in model configuration and data other than addition of stock assessment data for the period 2019-2021. The new data have revised F and SSB in the most recent years in the 2019 keyrun, with the largest changes for herring. This revision is also seen in the ICES single species assessment. Estimated values have changed slightly, with the largest for the M2 of 0-groups of herring and sprat.



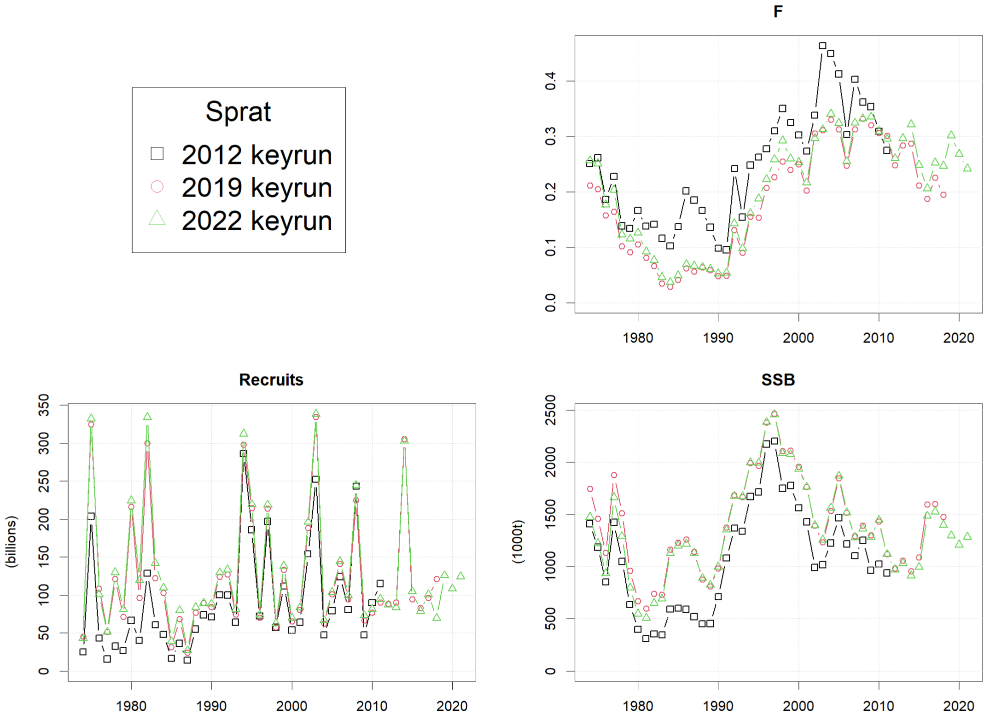


**Figure 5.2-15. Comparison of the 2012, 2019 and 2022 keyruns. Summary output.**

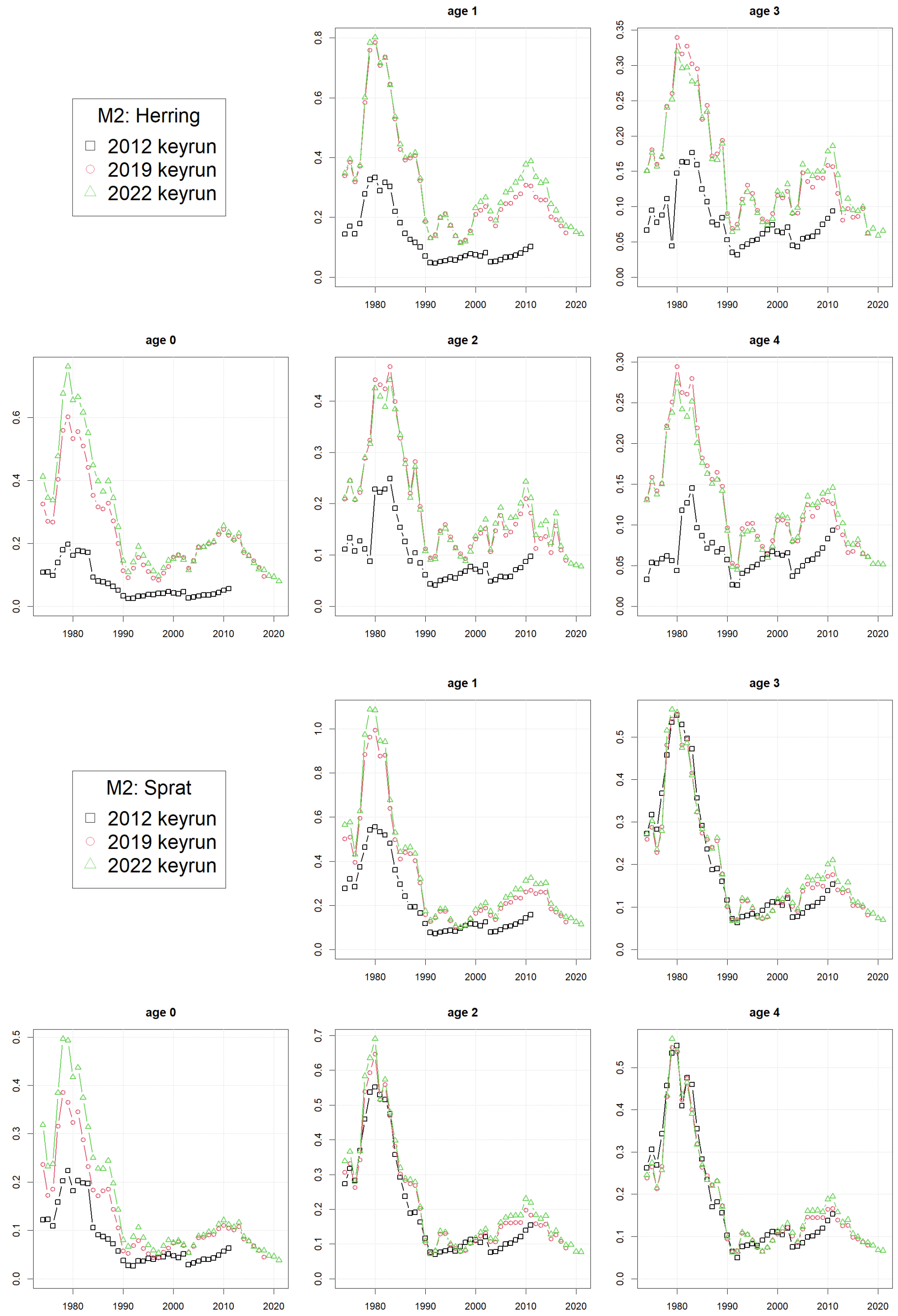


**Figure 5.2-16. Comparison of the 2012, 2019 and 2022 keyruns. Predation mortality (M2)**

### Comparison with the Gadget model run.

A comparison of the SMS keyrun and the results from the Gadget model evaluated at WGSAM 2019 is presented for herring (Figure 5.2-17) and for sprat (Figure 5.2-18). The estimated M values are quite similar.

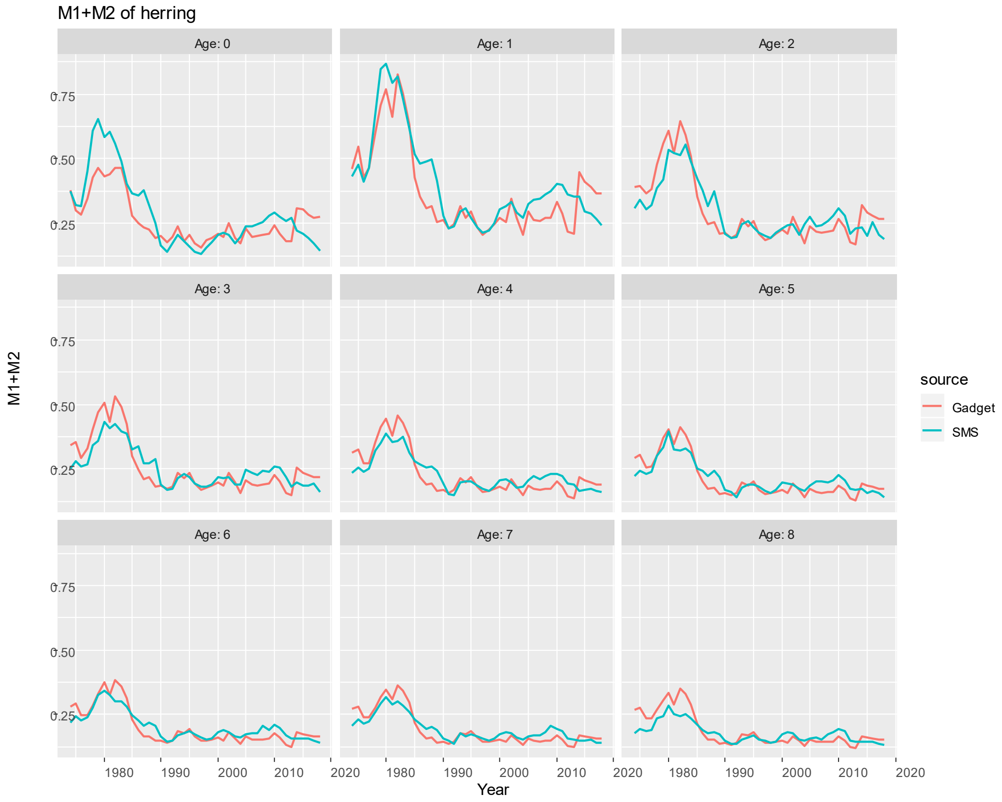


**Figure 5.2-17. Comparison of M (M=M1+M2) values for herring from the SMS 2019 keyrun and a Gadget configuration.**

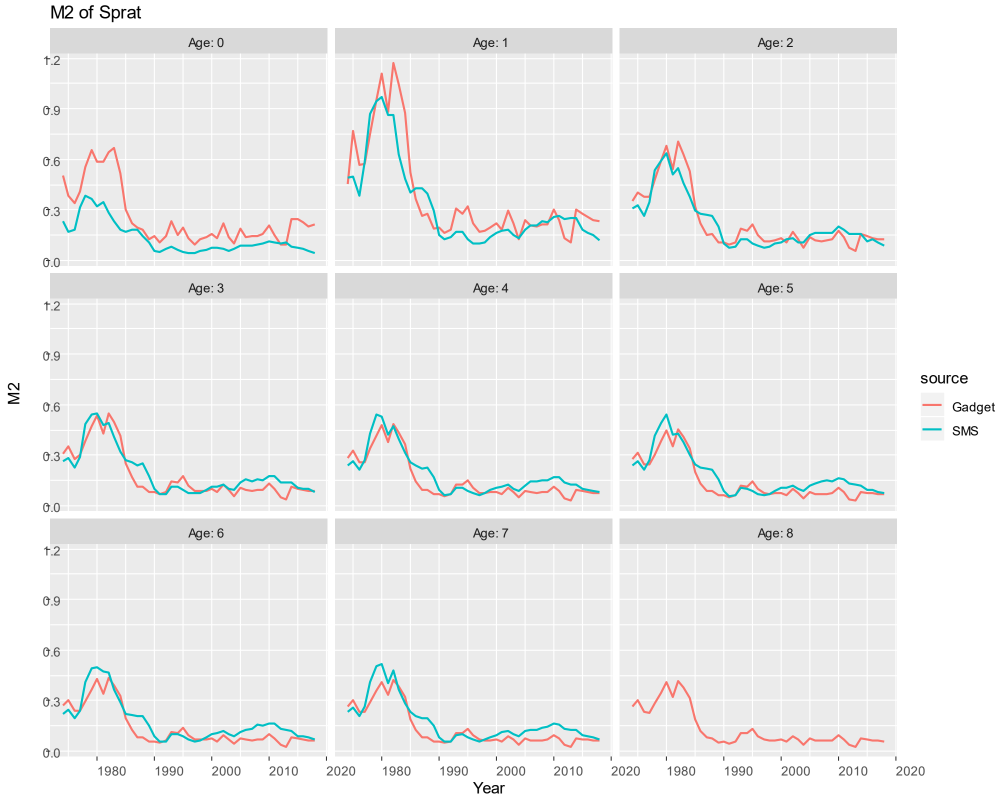


**Figure 5.2-18. Comparison of M (M=M1+M2) values for sprat from the 2019 keyrun and a Gadget configuration.**

## Conclusion, keyruns

### 2019 keyrun

WGSAM 2019 discussed and reviewed the changes in input data and the results in detail and concluded that:

The (2019) key-run as currently best possible run with SMS to provide natural mortality estimates. WGSAM recommends to use these values as input to single species stock assessments. The full time series should be used and not only an update for the years after the last key-run in 2012.

However, there are also clear limitations with the approach and results have been shown to be sensitive to e.g., consumption rates, assumptions regarding M1 and treatment of “Other Food” as well as the size selectivity of cod. In addition, the results depend to a large extent on the outcome of the ICES Eastern Baltic cod assessment. Any bias in this assessment directly influences the predation mortality estimates. Assumptions around other food and constant vulnerabilities may also bias the natural mortality estimates to some extent. Contrarily, the very similar results from the Gadget model run are encouraging and increase the credibility of the provided M time series.

WGSAM (2019) does not recommend using the uncertainty estimates around M as these are underestimated due to the assumption that the cod population is known without error.

#### 2022 keyrun

WGSAM 2022 reviewed the updated keyrun for the central Baltic Sea and concluded that it provides the best possible estimates of natural mortality for the Baltic sprat and central Baltic herring stocks. WGSAM recommends using these values as input to single species stock assessments. The full time series should be used and not only an update for the years after the last key-run in 2019.

Main aspects of uncertainty and limitations highlighted for the 2019 model remain also for this updated keyrun.

### Identified areas of priority research

#### WGSAM 2019 recommendation

WGSAM 2019 recommended:

1. More analyses on stomach data to get a better process understanding what is driving the systematic changes in relative stomach contents.

2. A split of Other Food in parts where the time dynamic can be taken into account (e.g., flatfish and *Saduria entomon*) and a part that still needs to be assumed constant in time may be beneficial.

3. The inclusion of spatial dynamics (either directly or via overlap coefficients) may improve the fit to data sources.

4. A run with age 1 as recruits could be tried because input for the 0 group is highly uncertain.

5. Account for the uncertainty in cod numbers in the model.

Some work has been done since 2019 to address these recommendation. The development of the R-package FishStomachs allows a consistent approach to analyse stomach data (**Recommendation 1**). Several papers (e.g. Neuenfeldt et al., 2022) discuss the poor state of the cod in the Eastern Baltic Sea in relation to e.g. food availability, parasite load and effects of environmental changes. A model like SMS is however not adequate to handle such issues.

With respect to **recommendation 5**) the SMS model was extended such that uncertainties of the stock size of “other predators” could be taken into account in e.g. the estimation of the uncertainties on M2 values. The approach uses “observations” of stock size (e.g. 10 samples of stock size drawn from the “real” (WGBFAS estimate) cod stock size with a specified CV. These “observations” are used in the SMS likelihood where the “observed” and “real” data are used to estimate a model stock size. This resulting uncertainty of the model stock size are afterwards used in the calculation of e.g. M2 values, in a similar way as for all other SMS output. The likelihood contributions from comparing the model stock size with the “real” stock size are used together with the likelihood contributions from the catch, cpue and stomach observations. Preliminary runs with the new version of SMS show that the best likelihood is obtained with a relative poor fit between the model cod stock size and the “real” stock size. The new SMS version has tendency to estimate a considerably lower stock size of cod than the “real” stock size. Said in another way, the likelihood for all (other) observations in SMS become better if the model stock size is decreased, on the cost of a poor fit between model stock size and "real" stock size. More work is needed, before the new version of SMS can be applied.

#### WGSAM 2022 recommendation

For further work on the Baltic Sea SMS keyrun, WGSAM 2022 recommends the following:

1. Additional analyses of the cod diet to get a better process understanding of what is driving the systematic changes in relative stomach contents.

2. A split of Other Food in parts where the time dynamic can be considered (e.g., flounder and *Saduria entomon*) directly or via a proxy and a part that still needs to be assumed constant in time may be beneficial. The availability of the different food groups for cod in the Baltic Sea have changed strongly over time (e.g. Haase et al. 2020).

In 2019, the WGSAM reviewers made these two recommendations, on which no work has been done since then. The overestimations of the model in the amount of herring eaten by cod and the underestimations of other food and sprat (all in recent years, Figure 8), indicate that it is still a good idea to consider these recommendations. The recent changes in the condition of Baltic cod indicate that there has been likely a change in the food availability for cod. Although it could be difficult to implement these changes they could be quite important for reliable estimates of the model.

3. The inclusion of spatial dynamics in the food consumption of cod (either directly or via overlap coefficients) may improve the fit to data sources.

In 2019, the WGSAM reviewers made this recommendation, which has not been considered so far. As long as the spatial overlap between prey and predator is constant through time, incorporating an overlap coefficient should not have a strong effect on the model predictions as this is currently addressed via the vulnerability parameter. However, if there is a temporal trend in the spatial overlap between predator and the prey species, and/or if the prey age classes are unevenly distributed through space, an overlap coefficient (per prey age class) could make a difference for the model predictions.

4. Uncertainty and potential issues with the quality of the estimates for the age 0 prey classes before they could be used as input in a single species stock assessment. Discussion with the stocks experts from the assessment working group on possible alternative data sources on the weight-at-age of age0 clupeids would be beneficial.

5. Account for the uncertainty in cod number at age in the model, such that this uncertainty is reflected in the uncertainties of estimated values like M2 and SSB.

In 2019, the WGSAM reviewers made this recommendation, which has not been fully considered so far. Due to the current problems regarding the age-reading of Baltic cod, only a length-based abundance estimate is available for this stock. Consequently, cod is considered as a predator with a constant density in the model. The effect of the uncertainties of cod stock densities and size distribution on the M2 predictions of SMS have not been tested.

6. Provide a multispecies $F_{MSY}$ estimation that can be used as a comparison with the $F_{MSY}$ values from the stock-by-stock (SS) frameworks currently used for advice.

The current model version of SMS of the Baltic Sea is not suitable for predictions of multispecies $F_{MSY}$ values. Unfortunately, due to the non-dynamic nature of cod in the

2022 key run, it is not straightforward to predict how fisheries mortality on cod may affect the other species in the food web through indirect effects. It is advisable to calculate this metric in the future if possible. The consideration of optimal yield trade-offs amongst the commercially exploited species in the complex is an important facet in improving fisheries management advice and an ecosystem-based approach to fisheries management.

# APPENDIX 1: SMS, a stochastic age–length structured multispecies model applied to North Sea and Baltic Sea stocks

### Working document to ICES WKMULTBAL, March 2012

By Morten Vinther and Peter Lewy,

DTU Aqua. Technical University of Denmark, National Institute of Aquatic Resources, Charlottenlund Castle, DK-2920 Charlottenlund, Denmark.

### Overview

SMS (Stochastic Multi Species model) is a fish stock assessment model which includes estimations of predation mortalities from observation of catches, survey indices and stomach contents. Estimation of predation mortality is based on the theory for predation mortality as defined by Andersen and Ursin (1977) and Gislason and Helgason (1985). SMS is a "forward running" model that operates with a chosen number of time steps (e.g. quarters of the year). The default SMS is a one-area model, but the model has options for spatial explicit predation mortality given a known stock distribution.

The model parameters are estimated using a maximum likelihood (ML) technique. Uncertainties of the model parameters are estimated from the Hessian matrix and confidence limits of derived quantities like historical fishing mortalities and stock abundances are estimated from the parameter estimates and the delta-method. SMS can be used to forecast scenarios and Management Strategy Evaluations, where fishing mortalities are estimated dynamically from Harvest Control Rules.

This document describes the model structure and the statistical models used for parameter estimation.

### Model Structure

#### Survival of the stocks

The survival of the stocks is described by the standard exponential decay equation of stock numbers (N).

$$N_{s,a,y,q+1} = N_{s,a,y,q}\ e^{-Z_{s,a,y,q}} \qquad \text{Eq. 1}$$

or

$$N_{s,a+1,y,+1,q=1} = N_{s,a,y,q=last\ season}\ e^{-Z_{s,a,y,q=last\ season}} \qquad \text{Eq. 2}$$

The instantaneous rate of total mortality, $Z_{s,a,y,q}$ by species s, age group a, year y and season q, is divided into three components; predation mortality (M2), fixed residual natural mortality (M1) and fishing mortality (F):

$$Z_{s,a,y,,q} = M1_{s,a,q} + M2_{s,a,y,q} + F_{s,a,y,q}$$

For non-assessment species which act as predators (e.g. grey seal and horse mackerel) stock numbers are assumed known and must be given as input.

#### Fishing mortality

Fishing mortality, $F_{s,a,y,q}$ is modelled from an extended separable model including age, year and season effects. However, as these effects may change over time a more flexible structure is assumed, allowing for such changes for specified periods. For convenience, the species index is left out in the following:

$$F_{a,y,q} = F^1_{Y,A1}\, F^2_y\ F^3_{Y,A2,q} \qquad \text{Eq. 3}$$

where indices $A1$ and $A2$ are grouping of ages, (e.g. ages 1–3, 4–7 and 8–9) and $Y$ is grouping of years (e.g. 1975–1989, 1990–2011).

Eq. 3 defines that the years included in the model can be grouped into a number of period clusters ($Y$), in which the age selection ($F^1$) and seasonal selection ($F^3$) are assumed constant. $F^2$is the year effect, specifying the overall level of F for a particular year. The grouping of ages for age selection, $A1$, and season selection, $A2$, can be defined independently.

2.2.1 Options for year effect

Given a good relationship between F and effort the fishing mortality can be calculated from the observed effort.

$$F_{a,y,q} = F^1_{Y,A1}\ EFFORT_y\ F^1_{Y,A2,q}$$

#### Natural Mortality

Natural mortality is divided into two components, predation mortality (M2) caused by the predators included in the model and a residual natural mortality (M1), which is assumed to be known and is given as input.

M2 of a prey species, $prey$, with size group $l_{prey}$ due to a predator species, $pred$, with size group $l_{pred}$ is calculated as suggested by Andersen and Ursin (1977) and Gislason and Helgason (1985).

$$M2_{prey,l_{prey},y,q} = \sum_{pred}\sum_{l_{pred}} \frac{\bar{N}_{pred,l_{pred},y,a}\ \ RA_{pred,l_{pred},y,q}\ \ S_{prey,pred,q}(l_{prey}, l_{pred})}{AB_{pred,l_{pred},y,a}} \qquad \text{Eq. 4}$$

where $RA$ denotes the total food ration (weight) of one individual predator per time unit, where S denotes the food suitability defined in section 6.2.3.2 and where AB is the total available (suitable) biomass. AB is defined as the sum of the biomass of preys weighted by their suitability. This total prey biomass includes also the so-called "other food" (OF) which includes all prey items not explicitly modelled, e.g. species of invertebrates and non-commercial fish species. Other food species are combined into one group, such that the total available prey biomass becomes:

$$AB_{pred,l_{pred},y,q} = \sum_{prey}\sum_{l_{prey}} \left(\bar{N}_{prey,l_{prey},y,q}\ W_{prey,l_{prey},y,q}\ S_{prey,pred,q}\left(l_{prey}, l_{pred}\right)\right) + OF_{pred,}\ S_{OF,pred,q}(l_{pred})$$ Eq. 5

M2 cannot directly be calculated from Eq. 4 because M2 also is included in the right hand term of Eq. 6 to calculate $\bar{N}$.

$$\bar{N} = \frac{N\,(1 - e^{-(M1+M2+F)})}{M1 + M2 + F}$$ Eq. 6

As no analytical solution for $M2$ exists, $M2$ has to be solved numerically. If the time step considered is sufficiently small, for instance a quarter, $M2$ becomes small and can optionally be approximated by replacing the average number during the season, $\bar{N}$, on the right hand side of Eq. 4 by the stock at the beginning of the season, N. As the right hand side of equation now is independent of M2 this quantity can be calculated directly from Eq. 4 where AB (Eq. 5) is modified correspondingly.

#### Use of size distribution by age

The equations outlined in the section above provide M2 at-size groups. However, predation mortality by age is needed as well because F and catches are age-structured. If just one size group per age group of predators and preys is assumed Eq. 4 can be used directly where the age index substitutes the size group index in stock numbers ($\bar{N}_{prey,a,y,q}\ =\ \bar{N}_{prey,l_{prey},y,q}$)

Given more size groups per age, the calculation of M2 at-age requires age–length-keys to split N at age to N at size group.

$$N_{s,l_s,y,q} = \sum_{a} N_{s,a,y,q}\ ALK_{s,a,l_s,y,q}$$ Eq. 7

where $ALK_{s,l_s,a,y,q}$ denotes the observed proportion of size group ls for a given species and age group, i.e. $\sum_{l_s} ALK_{s,l_s,a,y,q} = 1$

Assuming that F and M1 depend only on age and that M2 only depends on length, M2 at-age is estimated by: (leaving out the species, year and quarter indices).

$$\begin{aligned} M2_a &= Z_a \frac{\sum_l \bar{N}_{a,l}\, M2_{a,l}}{D_a} \\ &= \log(\frac{N_a}{N_a - D_a})\ \frac{\sum_l \bar{N}_{a,l}\ M2_l}{D_a} \end{aligned}$$

where

$$\begin{aligned} \bar{N}_{a,l} &= N_{a,l}\ \frac{1 - e^{-(F_{a,l}+M1_{a,l}+M2_{a,l})}}{F_{a,l} + M1_{a,l} + M2_{a,l}} \\ &= N_{a,l}\ \frac{1 - e^{-(F_a+M1_a+M2_l)}}{F_a + M1_a + M2_l} \end{aligned}$$

and where

$$D_a = \sum_l \bar{N}_{a,l} \, (F_a + M1_a + M2_l)$$

denotes the number of individuals at-age that died within a season.

### Food suitability

As suggested by Andersen and Ursin (1977) and Gislason and Helgason (1985) the size-dependent food suitability of prey entity $j$ for predator entity $i$ is defined as the product of a species dependent vulnerability coefficient, $\rho_{i,j}$, a size preference coefficient $\varrho_{i,j}(l_i, l_j)$, and an overlap index $o_{i,j,q}$. Suitability is then defined as:

$$S_{pred,prey,q}\left(l_{pred}, l_{prey}\right) = \rho_{pred,prey} \;\; \varrho_{pred,prey}\left(l_{pred}, l_{prey}\right) \;\; o_{pred,prey,q} \qquad \text{Eq. 8}$$

For the "other food" part suitability is defined as:

$$S_{OF,pred,q}\left(l_{pred}\right) = \rho_{OF,pred} \;\; o_{OF,pred,q} \; \exp\left(\upsilon_{pred} \; \log\left(W_{pred,l_{pred},q}/\overline{W}_{pred}\right)\right) \qquad \text{Eq. 9}$$

Where $\overline{W}_{pred}$ is the average size of the predator species. Eq. 9 extends the original equation, to allow predator size dependent suitability for other food, for values of $\upsilon_{pred}$ different from zero. The overlap index may change between seasons, but is assumed independent of year and sizes.

### Log-normal distributed size selection

Several functions can be used for the size preference for a prey. Andersen and Ursin (1977) assumed that a predator has a preferred prey size ratio and that a prey twice as big as the preferred size is as attractive as another half the prey size. This was formulated as a log-normal distribution:

$$\varrho_{pred,prey}\left(l_{pred}, l_{prey}\right) = \exp\left(-\frac{\left(\log\left(\frac{W_{l_{pred}}}{W_{l_{prey}}}\right) - \eta_{PREF\,pred}\right)^2}{2\,\sigma^2_{PREF\,pred}}\right); 0 < \varrho \leq 1 \qquad \text{Eq. 10}$$

Where $\eta_{PREF}$ is the natural logarithm of the preferred size ratio, $\sigma^2_{PREF}$is the "variance" of relative preferred size ratio, expressing how selective a predator is with respect to the size of a prey and where $W_{l_s}$is the mean weight for a species size group.

The basic size selection equation (Eq. 10) has been extended by modifying the preferred size ratio parameter.

$$\varrho_{pred,prey}\left(l_{pred}, l_{prey}\right) = \exp\left(-\frac{\left(\log\left(\frac{W_{l_{pred}}}{W_{l_{prey}}}\right) - \left(\eta_{PREF\,pred} + \xi_{prey} + \varpi_{pred}\;\log\left(W_{l_{pred}}\right)\right)\right)^2}{2\,\sigma^2_{PREF\,pred}}\right) \qquad \text{Eq. 11}$$

Where $\xi_{prey}$ specify a prey-specific adjustment term for the preferred size ratio, and where $\varpi_{pred}$ specifies how the preferred size range can change by predator size.

### Uniform size selection

Alternatively, a uniform size preference can be assumed within the range of the observed size ratio and zero size selection outside that ratio:

$$\varrho_{pred,prey}\left(l_{prep}, l_{prey}\right) = \begin{cases} 1 \quad \text{for } \eta_{MIN_{pred,prey}} \leq \dfrac{W_{l_{pred}}}{W_{l_{prey}}} \leq \eta_{MIN_{pred,prey}} \\ 0 \quad \text{for values outside observed range} \end{cases} \qquad \text{Eq. 12}$$

where $\eta_{MIN}$ and $\eta_{MAX}$ are the observed minimum and maximum predator/prey size ratios.

#### *7.2.3.2.2.1. Constraint uniform size selection*

The uniform size preference does not take into account that the preferred predator/prey size ratio might change by size, such that larger individuals select relatively smaller preys (Floeter and Temming, 2005; Sharft *et al.*, 2000). A way to account for that is to assume that the fixed minimum and maximum constants, $\eta_{MIN}$ and $\eta_{MAX}$, depend on the predator size:

$$_{pred,prey}\left(l_{pred}, l_{prey}\right) = \begin{cases} 1 \ \text{ for } U1_{pred,prey} + U2_{pred,prey}\ \log(W_{l_{pred}}) \leq \ \log\left(\dfrac{W_{l_{pred}}}{W_{l_{prey}}}\right) \leq U3_{pred,prey} + U4_{pred,prey}\ \log(W_{l_{pred}}) \\ 0 \quad \text{for values outside regression range} \end{cases} \qquad \text{Eq. 13}$$

The regression parameters are estimated externally by quantile regression (e.g. Koenker and Bassett, 1978) using e.g. the 2.5% and 97.5% percentiles of stomach content data. Figure 7.1 shows an example of such regression.

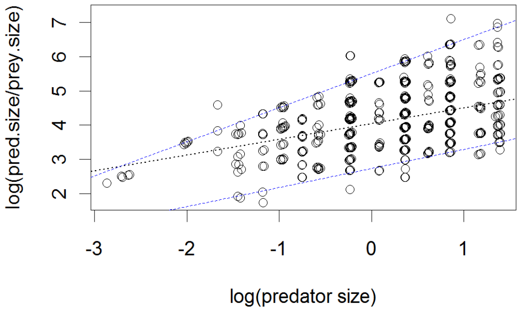


**Figure 7.1. Quantile regression of stomach contents observations (Baltic cod eating cod), with 2.5%, 50% and 97.5% lines shown. Predator and prey size in weight.**

### Adjustment of age–size keys

For the North Sea configuration, age length keys were obtained from the IBTS surveys where the same gear (i.e. the GOV trawl) has been used in the period considered. This allows an adjustment of the observed ALK's to account for mesh size selection. Using a logistic length-dependent selection function, selection is defined as:

$$SL_s(l) = 1/\left(1 + e^{(S1_s - S2_s * l)}\right)$$

Where $S1_s$ and $S2_s$ are species-specific gear selection parameters.

The adjusted ALK can then be derived from the observed ALK by:

$$ALK_{s,l_s,a,y,a} = \text{ObservedALK}_{s,l_s,a,y,q} \;/\; SL_{s,l_s}$$

which finally has to be standardised to 1 for each age before used in Eq. 7.

### Growth

Not implemented yet!

### Food ration

Food ration, RA, pr. time step is given as input or estimated from mean weight by size group assuming an exponential relationship between ration and body weight W.

$$RA_{pred,l_{pred},q} = \gamma_{pred,q} \; W_{pred,l_{pred}}^{\varsigma_{pred}} \qquad \text{Eq. 2}$$

where the coefficient $\gamma$ and ς are assumed to be known.

Body weight at-size group $lpred$ is estimated from mean length within the size group and a length–weight relationship.

### Area-based SMS

SMS has three area explicit options:

1 ) Default one area model. Both F and M2 are calculated for the entire stock area;

2 ) M2 by area. M2 is calculated by subareas, but F is assumed global;

3 ) M2 and F by area. Both M2 and F are calculated by area (forecast only).

#### Stock distribution

For the area-based models, the stock is assumed redistributed between areas between each seasonal time step.

$$N_{s,a,y,q}^{area} = N_{s,a,y,q} \;\; DIST_{s,a,y,q,area}$$

Where DIST is a stock distribution key that sums up to 1

$$\sum_{area} DIST_{s,a,y,q,area} = 1$$

The calculation of M2 for Option 1) is provided in the previous section.

The method for option 3) is very similar, but the calculations must be done by each subarea separately.

$$Z_a^{area} = F_a^{area} + M1_a^{area} + M2_a^{area}$$

where $M2^{area}$ is calculated as given in Eq. 4.

Option 2) is the hybrid, where F is global but M is calculated by area.

$$Z_a^{area} = F_a + M1_a^{area} + M2_a^{area}$$

$\bar{N}$ in an area is calculated in the usual way

$$\bar{N}_a^{area} = N_a^{area} \frac{1 - e^{-Z_a^{area}}}{Z_a^{area}}$$

The total number of individuals died due to predation mortality (DM2) then becomes:

$$DM2_a = \sum_{area} M2_a^{area} \bar{N}_a^{area} \qquad \text{Eq. 3}$$

M2 for the whole stock can be estimated from:

$$M2_a = \log\left(\frac{N_a}{N_a - D_a}\right) \frac{DM2_a}{D_a}$$

where

$$D_a = \sum_{area} DF_a^{area} + DM1_a^{area} + DM2_a^{area}$$

and $DF$ and $DM1$ represent the number that died due to fishery and residual mortality (M1) and are calculated in similar ways as specified for DM2 (Eq. 3).

#### Area based suitability parameters

For the "one area" SMS suitability is defined by Eq. 8.

The area-based version of suitability uses an area-specific vulnerability and overlap index, while the size preference ($\varrho$) is assumed independent of area.

$$S_{pred,prey,q}^{area} \left(l_{pred}, l_{prey}\right) = \rho_{pred,prey}^{area} \; \varrho_{pred,prey}\left(l_{pred}, l_{prey}\right) \; o_{pred,prey,q}^{area}$$

### Statistical models

Three types of observations are considered: Total international catch-at-age; survey abundance indices and relative stomach content. For each type, a stochastic model is formulated and the likelihood function is calculated. As the three types of observations are independent, the total log likelihood is the sum of the contributions from three types of observations. A stock–recruitment (penalty) function is added as a fourth contribution.

#### Catch-at-age

Catch-at-age observations are considered stochastic variables subject to sampling and process variation. The probability model for these observations is modelled along the lines described by Lewy and Nielsen (2003):

Catch-at-age is assumed to be lognormally distributed with log mean equal to log of the standard catch equation. The variance is assumed to depend on age and season and to be constant over years. To reduce the number of parameters, ages and seasons can be grouped, e.g. assuming the same variance for age 3 and age 4 in one or all seasons. Thus, the likelihood function, LCATCH, associated with the catches is:

$$L_{CATCH} = \prod_{s,a,y,q} \frac{1}{\sigma_{CATCH\ s,a,q}\sqrt{2\pi}}\ exp\left(-\frac{\left(\log(C_{s,a,y,q}) - E(\log(C_{s,a,y,q}))\right)^2}{2\ \sigma^2_{CATCH\ s,a,q}}\right) \qquad \text{Eq. 4}$$

Where

$$E(\log(C_{s,a,y,q}))\ = \log(F_{s,a,y,q}\ \bar{N}_{s,a,y,q})$$

Leaving out the constant term, the negative log-likelihood of catches then becomes:

$$\begin{aligned} l_{CATCH}\ &= -\log(L_{CATCH}) \\ &\propto \text{NOY} \sum_{s,a,q} \log(\sigma_{CATCH\ s,a,q}) \\ &+ \sum_{s,a,y,q} \left(\log(C_{s,a,y,q}) - E(\log(C_{s,a,y,q}))\right)^2 / 2\sigma^2_{CATCH\ s,a,q} \end{aligned} \qquad \text{Eq. 5}$$

Where $NOY$ is the number of years in the time-series.

#### Annual catches

Catch-at-age numbers by quarter have not been available for some of the demersal North Sea stocks in recent years. For use in the default SMS configuration of the North Sea, where quarterly time step is used, it is assumed that the seasonal distribution (the $F^3$ parameter in Eq. 3) is known and given as input. The likelihood function is modified to make use of the observed annual catches.

$$E(\log(C_{s,a,y}))\ = \log\left(\sum_q F_{s,a,y,q}\ \ \bar{N}_{s,a,y,q}\right)$$

$$L_{CATCH} = \prod_{s,a,y} \frac{1}{\sigma_{CATCH\ s,a}\sqrt{2\pi}}\ exp\left(-\frac{\left(\log(C_{s,a,y}) - E(\log(C_{s,a,y}))\right)^2}{2\ \sigma^2_{CATCH\ s,a}}\right) \qquad \text{Eq. 6}$$

#### Survey indices

Similarly to the catch observations, survey indices, $CPUE_{survey,s,a,y,q}$ are assumed to be log-normally distributed with mean:

$$E(\log(CPUE_{survey,s,a,y,q}))\ =\ \log(Q_{survey,a}\ \bar{N}_{SURVEY\ s,a,y,q}) \qquad \text{Eq. 7}$$

where Q denotes catchability by survey and $\bar{N}_{SURVEY}$ is mean stock number during the survey period. Catchability may depend on a single age or groups of ages. Similarly, the variance of log cpue, $\sigma^2_{SURVEY}$ may be estimated individually by age or by clusters of age groups. The negative log-likelihood is on the same form as Eq. 4.

$$\begin{aligned} l_{SURVEY} &= -\log(L_{SURVEY}) \\ &\propto NOY_{survey,s} \sum_{survey,s,a} \log(\sigma_{SURVEY\,survey,s,a}) \\ &+ \sum_{survey,s,a,y} \left(\log(CPUE_{survey,s,a,y}) - E(\log(CPUE_{survey,s,a,y}))\right)^2 / 2\sigma^2_{SURVEY\,s,a} \end{aligned}$$

Eq. 8

#### Stomach contents

The stomach contents observations, which are the basis for modelling predator food preference, consist of the average proportions by weight of the stomach content averaged over the stomach samples in the North Sea. The model observations, $STOM_{pred,l_{pred},prey,l_{prey},y,q}$, are given for combinations of prey and predator species and size classes. In the following, we use entity $i$ for a combination of predator species and predator size class (e.g. saithe 50–60 cm) and entity $j$ for the combination of prey species and prey size class eaten by entity $i$. Model observations are therefore defined as $STOM_{i,j,y,q}$.

$STOM_{i,j,y,q}$ is assumed to be a set of stochastic variables subject to sampling and process variations. For a given predator entity the observations across prey entities $i$ are continuous variables which sum to one. Thus, the probability distribution of the stomach observations for a given predator including all prey/length groups needs to be a multivariate distribution defined on the simplex. As far as the authors know, the Dirichlet distribution is the only distribution fulfilling this requirement. Leaving out the year and season index, the Dirichlet density function for a predator entity $i$ with $k$ observed diet proportions $STOM_{i,1}, \dots STOM_{i,k-1} > 0$ and the parameters $p_1, \dots, p_k > 0$ follows the probability density given by:

$$\begin{aligned} f_i = f(STOM_{i,1}, \dots, STOM_{i,k-1} \mid p_{i,1}, \dots, p_{i,k}) \\ = \frac{\Gamma(p_i)}{\prod_{j=1}^{k} \Gamma(p_{i,j})} \prod_{j=1}^{k} STOM_{i,j}^{p_{i,j}-1} \end{aligned}$$

Eq. 9

Where

$$STOM_{i,k} = 1 - \sum_{j=1}^{k-1} STOM_{i,j}$$

and

$$p_i = \sum_{j=1}^{k} p_{i,j}$$

The mean and variance of the observations in the Dirichlet distribution are equal to:

$$E(STOM_{i,j}) = \frac{p_{i,j}}{p_i}$$

$$Var(STOM_{i,j}) = \frac{E(STOM_{i,j})\left(1 - E(STOM_{i,j})\right)}{p_i + 1}$$

Eq. 10

The expected value of the stomach contents observations is modelled using the theory developed by Andersen and Ursin (1977):

$$E(STOM_{i,j}) = \frac{\bar{N}_j \; W_j \; S_{i,j}(l_i, l_j)}{\sum_j \left(\bar{N}_j \; W_j \; S_{i,j}(l_i, l_j)\right) \; + \; OF_i \, S_{OF,i}(l_i)} = \frac{p_{i,j}}{p_i} \qquad \text{Eq. 11}$$

where the food suitability function, $S$, is defined by Eq. 8 and Eq. 9. We use the same assumptions as for the calculation of M2 (Eq. 4), the small time steps used in the model allow for an approximation of $\bar{N}_j$ by $N_j$ in Eq. 11.

Regarding the variance of stomach content observations in unpublished analyses of the present authors of data from the North Sea stomach-sampling project 1991 (ICES, 1997), the authors indicate that the relationship between the variance and the mean of the stomach contents may be formulated in the following way:

$$Var(STOM_{i,j,y,q}) = \frac{E(STOM_{i,j,y,q}) \; \left(1 - E(STOM_{i,j,y,q})\right)}{V_{pred} \;\; U_{i,y,q}} \qquad \text{Eq. 12}$$

where $U_{i,y,q}$ is a known quantity reflecting the sampling level of a predator entity, e.g. the number of hauls containing stomach samples of a given predator and size class. $V_{pred}$ is a predator species-dependent parameter linking the sampling level and variance. Equating Eq. 10 and Eq. 12 implies that:

$$P_{i,y,q} \; = \; V_{pred} \, U_{i,y,q} - 1. \qquad \text{Eq. 13}$$

Insertion of Eq. 13 into Eq. 11 results in:

$$P_{i,j,y,q} = \left(V_{pred} \; U_{i,y,q} \; - 1\right) \frac{\bar{N}_j \; W_j \;\; S_{i,j}(l_i, l_j)}{\sum_j \left(\bar{N}_j \;\; W_j \; S_{i,j}(l_i, l_j)\right) \; + \; OF_i \, S_{OF,i}(l_i)}.$$

The parameters, $p_{i,j,y,q}$ are uniquely determined through stock numbers, total mortality, suitability parameters and $V_{pred}$.

Assuming that the diet observations for the predator length groups are independent, the negative log likelihood function including all predators and length groups is derived from Eq. 9:

$$l_{STOM} \; = \; -\log(L_{STOM}) = -\sum_{i,j,y,q} \log(f_{i,j,y,q}) \qquad \text{Eq. 14}$$

#### Modification of the stomach contents model

The stomach contents observations, $STOM_{prey,l_{prey},pred,l_{pred},y,q}$ are given for combinations of prey and predator species and size classes. For a diet consisting of a large proportion "other food" and several species and prey size classes, the proportion of the individual combination of species and size becomes small (less than 0.1%) for several prey entities. Very small proportions, in combination with a modest sampling size per stratum, make the estimation of parameters impossible in some cases. To overcome the problem, SMS has an option to let the likelihood use proportion summed over all size classes for a given prey species such that the prey entity equals the species.

The same grouping of all sizes from a prey is applied when the uniform size selection option (Eq. 12 and Eq. 1) is used. The likelihood function is the same as used for stomach observations that include prey size.

### Stock–recruitment

In order to enable estimation of recruitment in the last year for cases where survey indices catch from the recruitment age is missing (e.g. saithe), and to estimate parameters for forecast use, a stock–recruitment relationship $R_{s,y} = R\left(SSB_{s,y} \mid \alpha_s, \beta_s\right)$ penalty function is included in the likelihood function.

Recruitment to the model takes place in the same season ($recq$) and at the same age ($fa$) for all species. It is estimated from the Spawning–Stock Biomass (SSB) in the first season ($fq$) of the year, and a stock–recruitment relation. SSB is calculated from stock numbers, proportion mature (PM) and mean weight in the sea.

$$SSB_{s,y} = \sum_{a} N_{s,y,a,q=recq} \; PM_{s,y,a,q=recq} \; W_{s,y,a,q=recq} \qquad \text{Eq. 15}$$

At present, the Ricker (Eq. 16), the Beverton and Holt (Eq. 17), segmented regression (Eq. 18) and geometric mean are implemented.

$$R_{s,y} = \alpha_s \, SSB_{s,y-fa,fq} \; e^{\left(\beta_s \, SSB_{s,y-fa,fq}\right)} \qquad \text{Eq. 16}$$

$$R_{s,y} = \frac{\alpha_s \; SSB_{s,y-fa,fq}}{1 + \beta_s \; SSB_{s,y-fa,q}} \qquad \text{Eq. 17}$$

$$R_{s,y} = \begin{cases} \alpha_s \, SSB_{s,y-fa,fq} & \text{for } SSB_{s,y-fa,fq} < \beta_s \\ \alpha_s \; \beta_s & \text{for } SSB_{s,y-fa,fq} < \beta_s \end{cases} \qquad \text{Eq. 18}$$

Assuming that recruitment is lognormally distributed, the negative log likelihood, $l_{SR}$ , equals:

$$\begin{aligned} &l_{SR} \\ &= -\log(L_{SR}) \\ &\propto NOY \sum_{s} \log(\sigma_{SR\ a}) \\ &+ \sum_{s,a,y} \left(\log\left(N_{ss,a=fa,y,q=recq}\right) - E\left(\log\left(R_{s,y}\right)\right)\right)^2 / 2\sigma_{SR\ s}^2 \end{aligned} \qquad \text{Eq. 19}$$

Where $NOY$ gives the number of years selected and where Eq. 20 gives the expected recruitment for the Ricker case.

$$E(\log(R_s)) = \log\left(\alpha_s\, SSB_{s,y-fa,fq}\, e^{(\beta_s\, SSB_{s,y-fa,fq})}\right) \qquad \text{Eq. 20}$$

### Total likelihood function and parameterisation

The total negative log likelihood function, $l_{TOTAL}$, is found as the sum of the four terms:

$$l_{TOTAL} = l_{CATCH} + l_{SURVEY} + l_{STOM} + l_{SR}$$

To ensure uniquely determined parameters it is necessary to fix part of them. For the F-at-age model (Eq. 3) the year selection in the beginning of each year range (Y) has been fixed to one ($F^2_{y=\text{first year in each group of years}} = 1$). The season effect in the last season of all years and ages is also fixed ($F^3_{y,a,q=last\ season} = 1/\text{number of seasons}$).

Eq. 4 and Eq. 8 indicate that it is only possible to determine relative vulnerability parameters, $\rho_{pred,prey}$. We have chosen to fix the vulnerability of other food for all predators to 1.0. Similarly, the biomass of other food OFpred has arbitrarily been set (e.g. at 1 million tonnes) for each predator. The actual value by predator was chosen to obtain estimates of vulnerability parameters for the fish prey at around 1. Other parameters than suitability are practically unaffected of the actual choice of biomass of other food.

In the food suitability function (Eq. 8 and Eq. 9), vulnerability and overlap effects cannot be distinguished. Hence the overlap parameters were fixed for at least one season. In practice, however, several combinations of overlap have to be fixed (at e.g. 1).

Initial stock size, i.e. the stock numbers in the first year and recruitment over years are used as parameters in the model while the remaining stock sizes are considered as functions of the parameters determined by Eq. 1 and Eq. 2.

The year effect ($F^2_{y,s}$) in the separable model for fishery mortality (Eq. 3) takes one parameter per species for each year in the time-series which sum up to a considerable number of parameters. To reduce this high number of parameters, the year effect can optionally be modelled from a cubic spline function which requires fewer parameters. The number of knots must be specified if this option is used.

Another way to reduce the number of parameters is to substitute the parameters $\sigma_{CATCH}, \sigma_{SURVEY}$ and $\sigma_{SR}$ used in the likelihood functions by their empirical estimates. This optional substitution has practically no effect on the model output and the associated uncertainty.

Appendage 1 gives an overview of parameters and variables in the model.

The parameters are estimated using maximum likelihood (ML) i.e. by minimizing the negative log likelihood, $l_{TOTAL}$. The variance/covariance matrix is approximated by the inverse Hessian matrix. Uncertainties of functions of the estimated parameters (such as biomass and mean fishing mortality) are calculated using the delta method.

### SMS forecast

SMS is a forward-running model and can as such easily be used for forecast scenarios and Management Strategy Evaluation (MSE). SMS used the estimated parameters to calculate the initial stock numbers and exploitation pattern used in the forecast. Exploitation pattern is assumed constant in the forecast period, but is scaled to a specified average F, derived dynamically from Harvest Control Rules (HCR). Recruits are produced from the stock–recruitment relation, input parameters and a noise term.

#### Recruitment

Recruitment is estimated from the available stock–recruitment relationships, f(SSB), (see Section 6.3.4) and optionally a lognormal distributed noise term with standard deviation std.

$$R = f(SSB)\ e^{(std\ NORM(0,1))} \qquad \text{Eq. 21}$$

Where NORM(0,1) is a random number drawn from a normal distribution with mean=0 and standard deviation 1. A default value for std can be obtained from the estimated variance of stock–recruitment relationship, $\sigma^2_{SR_s}$ (Eq. 19)

Application of the noise function for the lognormally distributed recruitment gives on average a median recruitment as specified by f(SSB). Optionally, recruitment can be adjusted with half of the variance, to obtain, on average, a mean recruitment given by f(SSB).

$$R = f(SSB)\ e^{(std\ NORM(0,1))}\ e^{\left(-(std^2/2)\right)} \qquad \text{Eq. 22}$$

#### Harvest Control Rules

Several HCRs have been implemented, e.g. constant F and the ICES interpretation of management according to MSY for both short- and long-lived species. Selected, more complex management plans in force for the North Sea and Baltic Sea species have also been implemented.

### Model validation

Model validation (in the years 2004–2009), was focused on the performance of the model using simulated data from an independent model and simulated data produced by the SMS model itself. The independent model was implemented using R (R Development Core Team. 2011) and include a medium complex North Sea configuration (nine species, of which four are predators and eight species preys). The simulation model follows the SMS model specification with an addition of von Bertallanfy growth curves to model mean length-at-age. Variance around mean length-at-age was assumed to increase by increasing age. This combined age–length approach made it possible to simulate all the data needed for model verification. Test dataset from the simulation model included 20 years of catch data, one survey time-series per species covering all years and ages, and four quarterly stomach samples in year ten including stomach observations for all predator length groups. Data from the independent simulation model was used to verify that the SMS model actually works as intended and to investigate model sensitivity with respect to observation errors on catch, survey cpue and stomach data.

To test if model parameters were identifiable when uncertainties estimated from real data were applied, the SMS model was modified to produce observations with the estimated observation noise of catch, survey and stomach data. The experiment consists of the following steps:

1 ) Estimate model parameters using the SMS model and available North Sea data.

2 ) Generate 100 set of input data from SMS output (expected catch numbers, survey indices and stomach observations) and their associated variance of these values).

3 ) Let SMS estimate 100 sets of parameters from the 100 sets of input data.

This procedure results in one set of "true parameters", $\theta = (\theta_1, \dots, \theta_k)$ and 100 sets of estimated parameters, $\hat{\theta}_j = (\hat{\theta}_{1,j}, \dots, \hat{\theta}_{k,j}),\ j = 1, \dots, k$. Based on the 100 repetitions, and for each of the k parameters, the mean and the standard deviation of the mean $\bar{\hat{\theta}}_i$ and $\sigma_i$ and hence the 95% confidence limits was calculated. Finally, the proportion of the parameters was calculated for which $\theta_i$ lies in the 95% confidence interval of $\bar{\hat{\theta}}_i$.

The test showed that parameters are identifiable for most "real" North Sea configurations. For some species with relatively few diet observations, size selection parameters (Eq. 11) and the variance parameter (V) linking the stomach sampling level to the variance of Dirichlet distribution (Eq. 12 and Eq. 13), were outside the 95% confidence interval of $\bar{\hat{\theta}}_i$.

A more informal testing of the model has been done by simply using the model. SMS has been applied to produce the so called keyrun for both the species rich North Sea system (ten species with stock number estimation including seven prey species, and 16 species of "other predators") (ICES, WGSAM 2011) and the species poor Baltic Sea (cod, herring and sprat, one predator and three prey species) (WGSAM 2008, WKMAMPEL 2009). In addition, the model has been used in single-species mode for the ICES advice of blue whiting in the North East Atlantic (WGWIDE 2011) since 2005 and several sandeel stocks in the North Sea since 2009 (WGNSSK 2011). For MSE purposes, the model has been applied for sandeel and Norway pout in the North Sea (AGSANNOP 2007 ), blue whiting and pelagic stocks in the Baltic (WKMAMPEL 2009) in both single and multispecies mode.

SMS is essentially an extension of the statistical models normally used for single-species stock assessment. This allows the use the long list of available diagnostics tools, e.g. residuals plots, and retrospective analysis, developed for model testing of submodels for catch-at-age and survey indices. For stomach observations, however, fewer established methods are available. To apply reliable residual plots for stomach observations residuals need to be independent, which are not the case for the stomach contents model as the observations with respect to prey entity sum to one. Instead, we do the following: Let the predator entity, year and quarter be given and consider the stomach contents observations following the Dirichlet distribution:

$$STOM_r = (STOM_{r,1}, \dots, STOM_{r,k-1}) \sim Dir(p_{r,1}, \dots, p_{r,k})$$

Where $r$ equals the combined entity of predator entity, year and quarter and where $p_{r,j}, j = 1, \ldots, k$ represent the Dirichlet parameters estimated. Instead of considering the weight proportions, $STOM$, we consider absolute weights in the stomachs, $W_{r,j}$, $j = 1, \ldots, k$, where

$$STOM_{r,j} = \frac{W_{r,j}}{\sum_j W_{r,j}}$$

If we assume that $W_{r,j}$, $j = 1, \ldots, k$ are independent and follow gamma distributions with the same scale parameter, $\theta_r$, i.e.

$$W_{r,j} \sim \Gamma(p_{r,j}, \theta_r) \;\; j = 1, \ldots, k$$

it is well known that $STOM_r$ follows the Dirichlet distribution. We now assume that the opposite is the case (we have to prove that!) and hence assume that the absolute weights, $W_{r,j}$ are independent gamma distributed variables. We then transform these observations to obtain normal distributed residuals: Leaving out the indices, we get that $U = pgamma(W, p, \theta)$, where $pgamma$ is the distribution function of the gamma distribution, is uniformly distributed. To obtain normally distributed variables, $U$ is finally transformed to $V = qnorm(U)$, where $qnorm$ equals the inverse of the distribution function of the standardized normal distribution. This means that $V$ is our new residuals for stomach contents observations.

To obtain the absolute weight of the prey entities form the relative stomach content, $STOM$, we have to know the total stomach weight for the predator entity. We have not extracted those from the basic observations, but simply assumed that the total weight in the stomach is proportional to the number of stomachs sampled for a given predator entity.

### Implementation

The SMS has been implemented using the AD Model Builder (Fournier *et al.*, 2011), which is freely available from ADMB Foundation (www.admb-project.org). ADMB is an efficient tool including automatic differentiation for Maximum likelihood estimation of many parameters in nonlinear models.

SMS configurations may contain more than 1000 parameters of which less than 5% are related to predation mortality. It is not possible to estimate all parameters simultaneously without sensible initial parameter values. Such values are obtained in three phases:

1 ) Estimate “single-species” stock numbers, fishing mortality and survey catchability parameters assuming that natural mortality (M1+M2) are fixed and known (i.e. as used by the ICES single-species assessments).

2 ) Fix all the “single-species” parameters estimated in step 1 and use the fixed stock numbers to estimate initial parameter values for the predation parameters.

3 ) Use the parameter values from step 1 and 2 as initial parameter values and re-estimate all parameters simultaneously in the full model including estimation of predation mortality M2.

Optimisation might potentially be dependent on the initial parameter values, however the same final result was obtained using the three steps above or using a configuration where step two is omitted. Using step two however, in general makes the estimation process more robust as extreme values and system crash are avoided.

## References

Andersen, K. P., and Ursin, E. 1977. A Multispecies Extension to the Beverton and Holt Theory of Fishing, with account of Phosphorus Circulation and Primary Production. Meddr. Danm. Fisk.- og Havunders. 7 319–435.

Fournier, D. A., H. J. Skaug, J. Ancheta, J. Ianelli, A. Magnusson, M. N. Maunder, A. Nielsen, and J. Sibert. 2011. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods & Software. doi: 10.1080/10556788.2011.597854.

Gislason, H., and Helgason, T. 1985. Species interaction in assessment of fish stocks with special application to the North Sea. Dana 5: 1–44.

ICES. AGSANNOP. 2007. Report of the *ad hoc* Group on Sandeel and Norway Pout (AGSANNOP). ICES CM 2007/ACFM:40. 62 pp.

ICES. SGMSNS. 2005. Report of the Study Group on Multi Species Assessment in the North Sea. ICES C.M. 2005/ D:06. 159 pp.

ICES. WGNSSK. 2006. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK). ICES CM 2006/ACFM:35.

ICES. WGNSSK. 2011. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK). ICES CM 2011/ACOM:13. 1197 pp.

ICES. WGSAM. 2008. Report of the Working Group on Multispecies Assessment Methods (WGSAM), ICES CM 2008/RMC:06. 107 pp.

ICES WGSAM. 2011. Report of the Working Group on Multispecies Assessment Methods (WGSAM), ICES CM 2011/ SSGSUE:10. 229 pp.

ICES. WKMAMPEL. 2009. Report of the Workshop on Multi-annual management of Pelagic Fish Stocks in the Baltic. ICES CM 2009/ACOM:38. 120 pp.

Koenker, R., and Bassett, G. 1978. Regression Quantiles. Econometrica 46:1 33–50.

Lewy, P., and Nielsen, A. 2003. Modelling stochastic fish stock dynamics using Markov Chain Monte Carlo. ICES J. Mar. Sci., 60: 743–752.

Nielsen, A., and Lewy, P. 2002. Comparison of the frequentist properties of Bayes and the maximum likelihood estimators in an age-structured fish stock assessment model. Can. J. Fish. Aquat. Sci. 59: 136–143.

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

## Appendage 1. Notation, parameters and variables

### Indices

| | |
|---|---|
| *a* | age |
| *area* | area with specific predation mortality |
| *A1, A2* | group of ages |
| *Fa* | first age group in the model |
| *i* | prey entity, combination of prey species and prey size group |
| *j* | predator entity, combination of predator group and predator size group |
| *l* | species size class |
| *lpred* | predator size class |
| *lprey* | prey size class |
| *other* | other food "species" |
| *pred* | predator species |
| *prey* | prey species |
| *q* | season of the year, e.g. quarter |
| *recq* | recruitment season |
| *s* | species |
| *survey* | survey identifer |
| *y* | year |
| *Y* | group of years |

### Parameters and variables

| | |
|---|---|
| *AB* | available (suitable) prey biomass for a predator |
| *ALK* | proportion at-size for a given age group. Input |
| *C* | catch in numbers. Observations |
| *Cpue* | catch in numbers per unit of effort. Observations |
| *D* | number died |
| *DM1* | number died due to M1 |
| *DM2* | number died due to M2 |
| *DF* | number died due to F |
| *F* | instantaneous rate of fishing mortality |
| $F^1$ | age effect in separable model for fishing mortality. Estimated parameter |
| $F^2$ | year effect in separable model for fishing mortality. Estimated parameter |
| $F^3$ | season effect in separable model for fishing mortality. Estimated parameter |
| *M1* | instantaneous rate of residual natural mortality. Input |
| *M2* | instantaneous rate of predation mortality estimated in the model |
| *N* | stock number |
| *Ns,a,y=first year,q=1* | Stock number in the first year of the model. Estimated parameters |
| *Ns,a=fa,q=recq* | Stock numbers at youngest age (recruitment). Estimated parameter |
| *OF* | Biomass of other food for a predator. Input |
| *Q* | catchability, proportion of the population caught by one effort unit. Estimated |
| *Rs,y* | recruitment calculated from stock–recruitment model |
| *RA* | food ration, biomass consumed by a predator. Input |
| *S* | suitability of a prey entity as food for a predator entity |
| *S1, S2* | mesh selection parameters. Estimated |
| *SSB* | spawning–stock biomass |
| *STOM* | weight proportion of prey i found in the stomach of predator j. Observations |
| *U* | sampling intensity of stomachs. Observation |

| | |
|---|---|
| $V$ | variance of diet observations in relation to sampling intensity. Estimated Parameter |
| $W$ | body weight. Input |
| $Z$ | instantaneous rate of total mortality |
| $\alpha$ | stock–recruitment parameter. Estimated |
| $\beta$ | stock–recruitment parameter. Estimated |
| $\varrho$ | prey size preference of a predator. Estimated parameter |
| $\gamma$ | food ration coefficients. Input |
| $\varsigma$ | food ration exponent. Input |
| $\upsilon$ | parameter for size dependent preference for other food. Estimated parameter |
| $\eta PREF$ | natural logarithm of the preferred predator prey size ratio. Estimated parameter |
| $\eta MIN$ | observed minimum relative prey size for a predator species. Input |
| $\eta MAX$ | observed maximum relative prey size for a predator species. Input |
| $o$ | spatial overlap between predator and prey species. Estimated parameter |
| $\rho$ | coefficient of species vulnerability. Estimated parameter |
| $\sigma CATCH$ | standard deviation of catch observations. Estimated parameter |
| $\sigma PREF$ | parameter expressing how particular a predator is about the size of its prey. Parameter |
| $\sigma SR$ | standard deviation of stock–recruitment estimate. Estimated parameter |
| $\sigma STOM$ | standard deviation of stomach content observations (used with lognormal distribution) |
| $\sigma SURVEY$ | standard deviation of survey cpue observations. Estimated parameter |

# APPENDIX 2: Option file for SMS-key-runs

Keyrun 2022

```
# sms.dat option file
# the character "#" is used as comment character, such that all text and numbers
# after # are skipped by the SMS program
#
########################################
# Produce test output (option test.output)
#  0 no test output
#  1 output file sms.dat and  file fleet.info.dat as read in
#  2 output all single species input files as read in
#  3 output all multi species input files as read in
#  4 output option overview
#
# 11 output between phases output
# 12 output iteration (obj function) output
# 13 output stomach parameters
# 19 Both 11, 12 and 13
#
# Forecast options
# 51 output hcr_option.dat file as read in
# 52 output prediction output summary
# 53 output prediction output detailed
0
########################################
# Produce output for SMS-OP program. 0=no, 1=yes
0
########################################
# Single/Multispecies mode (option VPA.mode)
# 0=single species mode
# 1=multi species mode, but Z=F+M (used for initial food suitability parm. est.)
# 2=multi species mode, Z=F+M1+M2
1
########################################
# Number of areas for multispecies run (default=1)
1
#&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&
#
# single species parameters
#
## first year of input data (option first.year)
1974
########################################
## first year used in the model (option first.year.model)
1974
########################################
## last year of input data (option last.year)
2021
########################################
## last year used in the model (option last.year.model)
2021
########################################
##  number of seasons (option last.season). Use 1 for annual data
4
########################################
## last season last year (option last.season.last.year). Use 1 for annual data
4
########################################
## number of species (option no.species)
3
########################################
# Species names, for information only. See file species_names.in
#  Cod Herring Sprat
########################################
## first age all species (option first.age)
0
########################################
## recruitment season (option rec.season). Use 1 for annual data
3
########################################
## maximum age for any species(max.age.all)
```

```
11
########################################
## various information by species
# 1. last age
# 2. first age where catch data are used (else F=0 assumed)
# 3. last age with age dependent fishing selection
# 4. Esimate F year effect from effort data. 0=no, 1=yes
# 5. Last age included in the catch at age likelihood (normally last age)
# 6. plus group, 0=no plus group, 1=plus group
# 7. predator species, 0=no, 1=VPA predator, 2=Other predator
# 8. prey species, 0=no, 1=yes
# 9. Stock Recruit relation
#      1=Ricker, 2=Beverton & Holt, 3=Geom mean,
#      4= Hockey stick, 5=hockey stick with smoother,
#      51=Ricker with estimated temp effect,
#      52=Ricker with known temp effect,
#      61=STN Ricker for sprat. Input from file Sprat_rec_61.in
#      71=STN special SSB/R for cod. Input from file Cod_rec_71.in
#      >100= hockey stick with known breakpoint (given as input)
# 10. Spawning season (not used yet, but set to 1)
# 11. Additional data for Stock Recruit relation
11 0 0 0 0 0 2 0 0 0 0   # 1 Cod as other predator
 8 1 5 0 8 1 0 1 3 0 0   # 2 Herring
 7 1 4 0 7 0 0 1 3 0 0   # 3 Sprat
########################################
## use input recruitment estimate (option use.known.rec)
#   0=estimate all recruitments
#   1=yes use input recruitment from file known_recruitment.in
0
########################################
## adjustment factor to bring the beta parameter close to one (option beta.cor)
      1e+06  #      Herring
      1e+06  #        Sprat
########################################
## year range for data included to fit the R-SSB relation (option
SSB.R.year.range)
# first (option SSB.R.year.first) and last (option SSB.R.year.last) year to
consider.
# the value -1 indicates the use of the first (and last) available year in time series
# first year by species
         -1  #      Herring
       1990  #        Sprat
# last year by species
         -1  #      Herring
         -1  #        Sprat
########################################
## Objective function weighting by species (option objective.function.weight)
# first=catch observations,
# second=CPUE observations,
# third=SSB/R relations
# fourth=stomach observations, weight proportions
# fifth=stomach observations, number at length
##
0 0 0    1 0  # 1 Cod
1 1 0.05 0 0  # 2 Herring
1 1 0.05 0 0  # 3 Sprat
########################################
## parameter estimation phases for single species parameters
# phase.rec (stock numbers, first age) (default=1)
1
# phase.rec.older (stock numbers, first year and all ages) (default=1)
1
# phase.F.y (year effect in F model) (default=1)
1
# phase.F.y.spline (year effect in F model, implemented as spline function)
-1
# phase.F.q (season effect in F model) (default=1)
1
# phase.F.a (age effect in F model) (default=1)
1
# phase.catchability (survey catchability) (default=1)
1
# phase.SSB.R.alfa (alfa parameter in SSB-recruitment relation) (default=1)
1
# phase.SSB.R.beta (beta parameter in SSB-recruitment relation) (default=1)
1
########################################
```

```
## minimum CV of catch observation used in ML-estimation (option min.catch.CV)
0.1
########################################
## minimum CV of catch SSB-recruitment relation used in ML-estimation (option min.SR.CV)
0.1
########################################
## Use proportion landed information in calculation of yield (option calc.discard)
#    0=all catches are included in yield
#    1=yield is calculated from proportion landed (file proportion_landed.in)
          0  #      Herring
          0  #        Sprat
########################################
## use seasonal or annual catches in the objective function (option combined.catches)
# do not change this options from default=0, without looking in the manual
#    0=annual catches with annual time steps or seasonal catches with seasonal
time steps
#     1=annual catches with seasonal time steps, read seasonal relative F from
file F_q_ini.in (default=0)
          0  #      Herring
          0  #        Sprat
########################################
## use seasonal or common combined variances for catch observation
# seasonal=0, common=1 (use 1 for annual data)
          1  #      Herring
          0  #        Sprat
########################################
##
# catch observations: number of separate catch variance groups by species
           2   #     Herring
           4   #       Sprat

#  first age group in each catch variance group
1 2    #  Herring
1 2 3 6        #  Sprat
########################################
##
# catch observations: number of separate catch seasonal component groups by species
           3   #     Herring
           2   #     Sprat

#  first ages in each seasonal component group by species
1 2 3   #  Herring
1 2     #  Sprat
########################################
## first and last age in calculation of average F by species (option avg.F.ages)
3 6  # Herring
3 5  # Sprat
########################################
## minimum 'observed' catch, (option min.catch). You cannot log zero catch at age!
#
# 0 ignore observation in likelihood
#
# negative value gives percentage (e.g. -10 ~ 10%) of average catch in age-group for input
catch=0
# negative value less than -100 substitute all catches by the option/100 /100 *average catch
in the age group for catches less than (average catch*-option/10000
#
# if option>0 then will zero catches be replaced by catch=option
#
# else if option<0 and option >-100 and catch=0 then catches will be replaced by catch=av-
erage(catch at age)*(-option)/100
# else if option<-100  and catch < average(catch at age)*(-option)/10000 then catches will
be replaced by catch=average(catch at age)*(-option)/10000
          0  #      Herring
          0  #        Sprat
########################################
# catch observations: number of year groups with the same age and seasonal selection
            3  #     Herring
            2  #       Sprat

#  first year in each group (please note #1 will always be changed to first model year)
1974 1989 2005 #  Herring
1974 2000  #  Sprat
########################################
##
# number of nodes for year effect Fishing mortality spline
# 1=no spline (use one Fy for each year), >1 number of nodes
           1   #     Herring
           1   #       Sprat
```

```
#  first year in each group
1974    #  Herring
1974    #  Sprat
########################################
## year season combinations with zero catch (F=0) (option zero.catch.year.season)
# 0=no, all year-seasons have catches,
# 1=yes there are year-season combinations with no catch.
#   Read from file zero_catch_seasons_ages.in
# default=0
0
########################################
## season age combinations with zero catch (F=0) (option zero.catch.season.ages)
# 0=no, all seasons have catches,
# 1=yes there are seasons with no catch. Read from file zero_catch_season_ages.in
# default=0
0
########################################
## Factor for fixing last season effect in F-model (default=1) (fix.F.factor))
          1  #      Herring
          1  #        Sprat
########################################
## Uncertainties for catch, CPUE and SSB-R observations (option calc.est.sigma)
#  values: 0=estimate sigma as a parameter (the right way of doing it)
#          1=Calculate sigma and truncate if lower limit is reached
#          2=Calculate sigma and use a penalty function to avoid lower limit
#  catch-observation, CPUE-obs, Stock/recruit
           0            0            0
########################################
# Read HCR_option file (option=read.HCR) default=0
#  0=no  1=yes
0
########################################
#
```

# # multispecies parameters

```
#
# Exclude year,season and predator combinations where stomach data are not incl.(option
incl.stom.all)
#   0=no, all stomach data are used in likelihood
#   1=yes there are combinations for which data are not included in the likelihood.
#      Read from file: incl_stom.in
#   default(0)
1
########################################
##  N in the beginning of the period or N bar for calculation of M2 (option use.Nbar)
#  0=use N in the beginning of the time step (default)
#  1=use N bar
0
########################################
## Maximum M2 iterations (option M2.iterations) in case of use.Nbar=1
5
########################################
## convergence criteria (option max.M2.sum2) in case of use.Nbar=1
#   use max.M2.sum2=0.0 and M2.iterations=7 (or another high number) to make
Hessian
0
########################################
## likelihood model for stomach content observations (option stom.likelihood)
#  1 =likelihood from prey weight proportions only (see option below)
#   2 =likelihood from prey weight proportions and from prey numbers to estimate size
selection
#  3 =Gamma distribution for prey absolute weight and size selection from prey numbers
1
########################################
# Variance used in likelihood model for stomach contents as prey weight propor-
tion (option stomach.variance)
#  0 =not relevant,
#  1 =log normal distribution,
#  2 =normal distribution,
#  3 =Dirichlet distribution
3
########################################
## Usage of age-length-keys for calc of M2 (option simple.ALK))
#  0=Use only one sizegroup per age (file lsea.in or west.in)
#  1=Use size distribution per age (file ALK_all.in)
0
########################################
```

```
## Usage of food-rations from input values or from size and regression parameters (option consum)
#  0=Use input values by age (file consum.in)
#  1=use weight at age (file west.in) and regression parameters (file consum_ab.in)
#  2=use length at age (file lsea.in), l-w relation and regression parameters (file consum_ab.in)
0
########################################
## Size selection model based on (option size.select.model)
#  1=length:
#      M2 calculation:
#         Size preference:
#           Predator length at age from file: lsea.in
#           Prey     length at age from file: lsea.in
#         Prey mean weight is weight in the sea from file: west.in
#      Likelihood:
#         Size preference:
#           Predator mean length per length group (file: stom_pred_length_at_sizecl.in)
#           Prey mean length per ength group (file stomlen_at_length.in
#              Prey mean weight from mean weight per prey length group (file: stomweight_at_length.in
#  2=weight:
#      M2 calculation:
#         Size preference:
#           Predator weight at age from file: west.in
#           Prey     weight at age from file: west.in
#         Prey mean weight is weight in the sea from file: west.in
#      Likelihood:
#         Size preference
#            Predator mean weight is based on mean length per predator length group (file: stom_pred_length_at_sizecl.in)
#              and l-w relation (file: length_weight_relations.in),
#           Prey mean weight per prey length group (file: stomweight_at_length.in)
#              Prey mean weight from mean weight per prey length group (file: stomweight_at_length.in
#  3=weight:
#       M2 calculation: Same as option 2
#       Likelihood:
#         Size preference:
#            Predator mean weight is based on mean length per predator length group (file: stom_pred_length_at_sizecl.in)
#              and l-w relation (file: length_weight_relations.in),
#             Prey mean weight per prey length group (file: stomlen_at_length.in) and l-w relation (file:length_weight_relations.in)
#           Prey mean weight from prey mean length per prey length group (file: stomlen_at_length.in) and l-w relation (file: length_weight_relations.in)
#  4=weight:
#       M2 calculation:
#         Size preference:
#             Predator mean weight from file lsea.in (length in the sea) and l-w relation (file: length_weight_relations.in)
#           Prey mean weight from file lsea.in (length in the sea) and l-w relation (file: length_weight_relations.in)
#       Likelihood:  Same as option 3
#  5=weight in combination with simple.ALK=1:
#       M2 calculation:
#         Size preference:
#            Predator weight based on length from file ALK_all.in (length distribution at age) and l-w relation (file: length_weight_relations.in)
#            Prey     weight based on length from file ALK_all.in (length distribution at age) and l-w relation (file: length_weight_relations.in)
#        Prey mean weight based on length from file ALK_all.in (length distribution at age) and l-w relation (file: length_weight_relations.in)
#       Likelihood: Same as for option 2
#  6=weight in combination with simple.ALK=1:
#       M2 calculation: Same as option 5
#       Likelihood: Same as option 3
2
########################################
# Adjust Length at Age distribution by a mesh selection function (option L50.mesh)
#  Please note that options simple.ALK shoud be 1 and option size.select.model should be 5
# L50 (mm) is optional given as input. Selection Range is estimated by the model
# L50= -1 do not adjust
# L50=0, estimate L50 and selection range
# L50>0, input L50 (mm) and estimate selection range
# by VPA species
          -1  #      Herring
          -1  #        Sprat
########################################
## spread of size selection (option size.selection)
#   0=no size selection, predator/preys size range defined from observations
#   1=normal distribution size selection
#   3=Gamma distribution size distribution
```

```
#    4=no size selection, but range defined by input min and max regression
parameters (file pred_prey_size_range_param.in)
#   5=Beta distributed size distribution, within observed size range
#   6=log-Beta size distributed, within observed size range
#
# by predator
          1 #          Cod
########################################
## sum stomach contents over prey size for use in likelihood for prey weight
proportions (option sum.stom.like)
#   0=no, use observations as they are; 1=yes, sum observed and predicted stomach
contents before used in likelihood for prey weight proportions
#
# by predator
          1  #          Cod
########################################
## # Use estimated scaling factor to link number of observation to variance for
stomach observation likelihood (option stom_obs_var)
#     0=no, do not estimate factor (assumed=1);  1=yes, estimate the factor;
2=equal weight (1) for all samples
#
# by predator
          1  #          Cod
########################################
## # Upper limit for Dirichlet sumP. A low value (e.g. 10) limits the risk of
overfitting. A high value (e.g. 100) allows a full fit. (option stom_max_sumP)
# by predator
        1000  #          Cod
########################################
## Scaling factor (to bring parameters close to one) for relation between no of
stomachs sampling and variance
#  value=0: use default values i.e. 1.00 for no size selection and otherwise 0.1
(option var.scale.stom)
          0  #          Cod
########################################
## other food suitability size dependency  (option size.other.food.suit)
#  0=no size dependency
#  1=yes, other food suitability is different for different size classes
          1  #          Cod
########################################
## Minimum observed relative stomach contents weight for inclusion in ML esti-
mation (option min.stom.cont)
        0.001  #          Cod
########################################
## Upper limit for no of samples used for calculation of stomach observation
variance (option max.stom.sampl)
     500 # 1e+06  #          Cod
########################################
##  Max prey size/ pred size factor for inclusion in M2 calc (option
max.prey.pred.size.fac)
        0.3  #          Cod
########################################
## inclusion of individual stomach contents observations in ML for weight pro-
portions (option stom.type.include)
# 1=Observed data
# 2= + (not observed) data within the observed size range (=fill in)
# 3= + (not observed) data outside an observed size range. One obs below and one
above (=tails)
# 4= + (not observed) data for the full size range of a prey species irrespective
of predator size (=expansion)
          1  #          Cod
########################################
## use overlap input values by year and season (use.overlap)
#   0: overlap assumed constant or estimated within the model
#    1: overlap index from file overlap.in (assessment only, use overlap from
last year in forecast)
#   2: overlap index from file overlap.in (assessment and forecast)
0
########################################
## parameter estimation phases for predation parameters
#  the number gives the phase, -1 means no estimation
#
#  vulnerability (default=2) (phase phase.vulnera)
2
# other food suitability slope (default=-1) (option phase.other.suit.slope)
2
# preferred size ratio (default=2) (option phase.pref.size.ratio)
```

```
2
# predator size ratio adjustment factor (default=-1) (option phase.pref.size.ratio.correction))
-1
# prey species size adjustment factor (default=-1) (option phase.prey.size.adjustment)
-1
# variance of preferred size ratio (default=2) (option phase.var.size.ratio)
2
# season overlap (default=-1) (option phase.season.overlap)
3
# Stomach variance parameter (default=2) (option phase.Stom.var)
2
# Mesh size selection of stomach age length key (default=-1) (option phase.mesh.adjust)
-1
########################################
```

# Annex 3: Multispecies Surplus Production Model

## Multi-Species Surplus Production Model (MSSPM) Georges Bank Keyrun

*Draft*

Howard Townsend (NMFS-OST), Ronald Klasky (NMFS-OST/ Caelum), Robert Gamble (NMFS-NEFSC)

October 10-14, 2022

ICES WGSAM - Woods Hole, MA

# Background

Surplus production models have been useful tools for estimating essential parameters for population dynamics and fisheries because they require minimal data and generate simple outputs (Prager, 1994; Hilborn and Walters, 1992). Though models with more forms can be used, the simplest of this type only require catch and biomass time series data, or their equivalents (Fox, 1975; Schnute, 1985; Ludwig and Walters, 1989; Prager, 1994), thus they can be broadly used for and make efficient use of data. Surplus production models, in some cases, have more accurately assessed stock than more complicated age-structured models (Ludwig and Walters, 1985).

Species interactions, such as predation and competition, may not have large discernible impacts on fisheries stocks, but may be an important feature in fishery ecosystems (Bax, 1998; Christensen, 1996). However, several authors (Sissenwine, 1984; Bax, 1991, 1998; Christensen, 1996) suggest that species interactions in combination with fisheries pressure may impact recruitment or suppress already diminished prey populations (Worm and Myers, 2003; Walters et al., 2005). A range of methods for incorporating multi-species interactions have been reviewed in Hollowed et al., 2000;Whipple et al., 2000; Plagányi, 2007).

With a move towards Ecosystem-Based Fishery Management (EBFM), a need to better account for interactions among living marine resources in an ecosystem arises. This is especially important because in many management plans there are mandates to achieve certain targets (e.g. MSY) that do not account for the effects of predators and prey, or competitors for common resources.  Numerous modeling approaches to account for ecological interactions have been developed and used over the past 30 years – EwE, Atlantis, OSMOSE, MSVPA, MSCAA; however, many of these approaches have drawbacks that limit their implementation for management. These models may be considered to be too complex or “data hungry”, so fisheries agencies may be reluctant to undertake multi-species modeling projects. This hesitance to attempt to quantify species interactions - a not at all trivial exercise - has continued to slow the uptake of EBFM. A key constraint in adopting EBFM has been the lack of modeling tools to easily implement these ecological considerations.

Multi-species surplus production models can be used to quantify species interactions, and because of their focus on minimal population structure, the barriers to begin modeling are surmountable due to lower data needs for each species than a corresponding age-structured assessment model. Generally, these include: biomass or survey index time series, catch data time series, spatial overlap and food habits information (for competitive and predation interactions). Multi-species models have been applied in a fisheries context since at least the 1970s. The background of multi-species (MS) models in general ecology goes back another 70 years. There have been major points of emphasis of these MS models in a fisheries context.

The progenitor of the approach we developed was a software package called Kraken (unpublished) which was developed initially as a simulation tool to explore the wide range of data we have in the Northeast Fisheries Science Center. The first model it incorporated was MS-PROD (Gamble & Link, 2009):

This model utilized spatial data and food habits data from the NEFSC bottom trawl surveys to develop reasonable parameters for competition and predation, bottom trawl survey indices to inform growth rates and carrying capacity to tune to the biological time series of commercially and recreationally important fish in the region. A second, simpler model was developed within Kraken, called AGG-PROD (Gamble & Link ,2012):

AGG-PROD used aggregate groups (groundfish, elasmobranchs and pelagics) instead of individual species. Both models showed an ability to be tuned to data, and to model direct and indirect effects based on changes to fishing pressure, ecological interactions and climate drivers. Kraken continued to be developed, and became more of a toolkit for creating multispecies surplus production models from simple Schaeffer type models with a single matrix that included competition, predation and density dependence, up to MS-PROD level models that also incorporated different functional forms for predation. Additionally, a genetic algorithm was explored as an estimation and fitting routine, but all work was used internally and some presented at conferences, but no results were published.

The current software and modeling approach, described next, drew from and greatly expanded on Kraken to create a flexible modeling and estimation package for surplus production models, complete with a more user friendly GUI.

The Multi-Species Surplus Production Model (MSSPM) software enables a user to build discrete-form single species and multi-species surplus production models. The software allows a user to 1) define a model, using choices of growth, harvest, competition, and predation forms, 2) run the model, which estimates the necessary parameters using choices for optimization algorithms, and 3) visualize the estimated data. A user can conduct diagnostic testing of a fitted model including retrospective analysis and parameter profiling. Forecasts may be generated using the estimated parameters and additional harvest and uncertainty parameters. The software has a specialized user interface, the REsource Management Options Review and Analysis (REMORA) tool that allows a non-technical user to explore fisheries management policy under different types and levels of uncertainty. More detail on the software is provided in Appendix A and in the User Guide (https://docs.google.com/document/d/14imgRumit05fcf3QJ00SXFoMu_837hvz/edit).

The MSSPM has been applied to data of the Georges Bank ecosystem. This application will allow further review of the model framework as well as a review of the potential utility of the MSSPM as a potential model to be used in a suite of models.

# Problem Identification

## Goal of the model

Generally, the goal of the MSSPM software is to allow an analyst to rapidly develop surplus production models of multiple fisheries species where some interactions are expected but data are limited. When data on stock age/length structure and diet habits are limited, but expected to be important, an analyst can use this approach to evaluate surplus production model structures (i.e., different population growth, species interaction, and harvest forms) and to test different hypotheses about an ecosystem. Upon selecting the best hypothesis model forms, a candidate set, the analyst can derive quantities of interest (e.g., estimated biomass, reference points) and produce forecasts for exploring policy options. In addition to being used for data-limited modeling, this approach can be useful for building a suite or ensemble of models with different structures to explore ecosystem-oriented fisheries management policies. Multi-model approaches are necessary for dealing with the uncertainty associated with model structure choices. Moreover, as this approach uses a simple model structure, the runtime can be fairly rapid, thus enabling a broad exploration of parameter uncertainty through profiling and forecasting.

## Use of the model

Though the software was built for building data-limited models, the application to Georges Bank is intended as a simple-structured model to be used with a suite of other models with a range of structures. This model can provide information about species interactions and environmental conditions that may be influencing the productivity of key Georges Bank species.

## Model decision support

When used with other multi-species ecosystem models, this model can be used to provide information on ecological reference points and to augment advice provided within a previously established framework.

As this is a biomass-based model, no age and length information are used. So as a standalone, this model application could not be used to provide advice where age/size limits are applied to catch, and would only be suitable for data-limited situations.

# Data to support the model

All data from this model application are found in the Georges Bank ms-keyrun data package (https://github.com/NOAA-EDAB/ms-keyrun/tree/master/data). Harvest data (catchIndex.rda) was used to drive the model dynamics. A combined survey index (surveyIndexAll.rda) was used for the relative biomass time series to fit the model to. A model-derived bottom temperature index was tested as a covariate to influence carrying capacity and population growth rates of some species. More information on this index is in Appendix B. Information to inform potential predation interactions was derived from "Fish Trophic Ecology of the Northeast U.S. Continental Shelf".

# Conceptual Model

# Available and alternative theories

The software gives flexibility when compared to the available theoretical population dynamics models. From a baseline discrete-form population model, an analyst can select different forms for population growth, harvest and predation. Alternative continuous-form and state-space-form population models have not been implemented in the software as of yet. Figure 1 provides a schematic of the software. Model equation terms can be selected for a flexible model structure (described in Appendix A "Creating a Population Model").

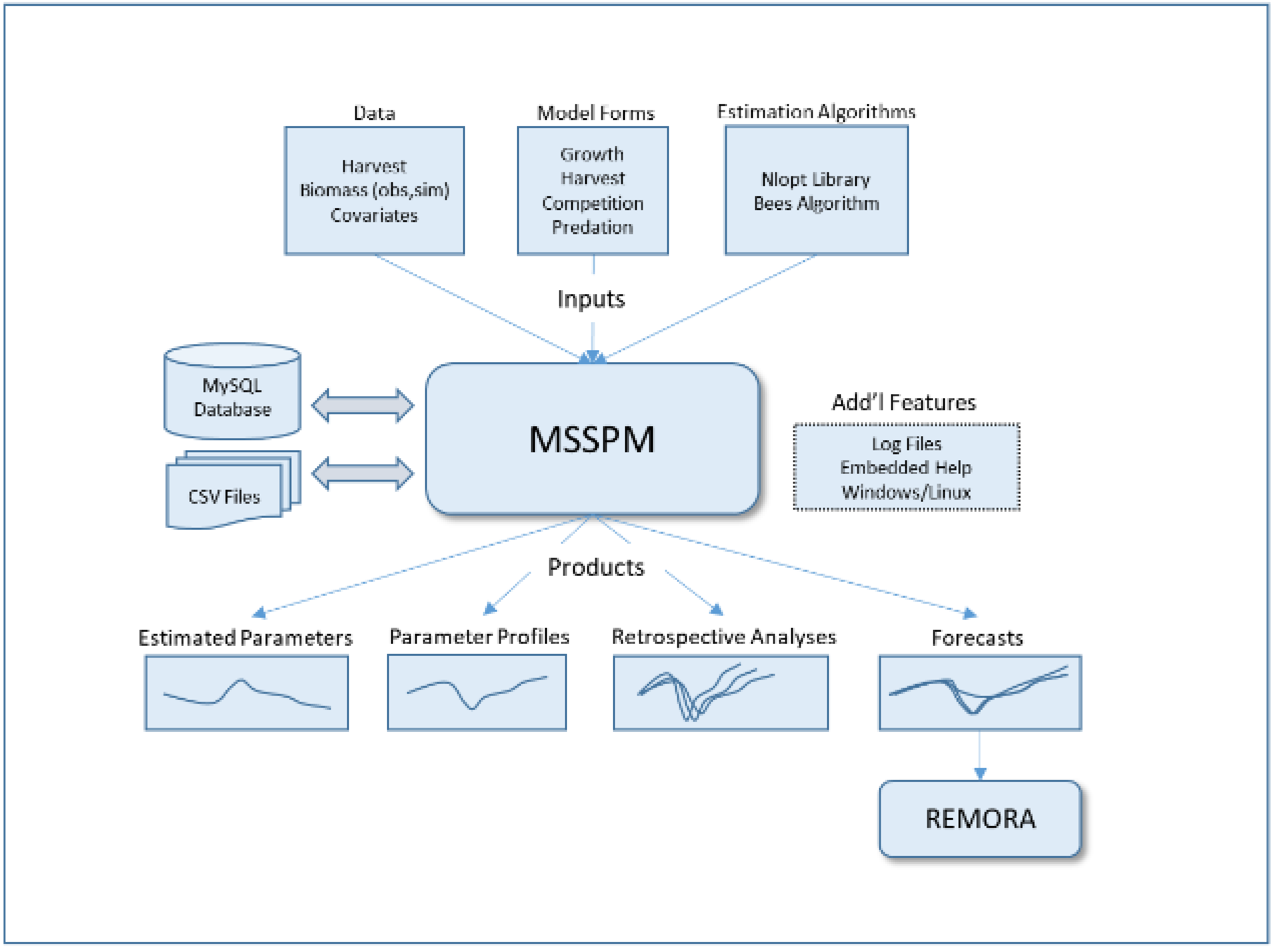


*Figure 1. Conceptual Model of MSSPM Software.*

# Assumptions

The main assumptions of surplus production models are that 1) all individuals have the same growth and mortality rates, 2) fish stocks, on the average, produce more offspring than necessary to replenish themselves, 3) populations are stable.

The software allows incorporation of environmental covariates that can be applied to population growth (r), carrying capacity (K), or survey catchability (Q). In this, a coefficient is estimated to determine the strength of the covariate effect. For example, the covariate is applied as:

$x = (1 + ccI)$ *Equation 1*

where, x is growth rate or carrying capacity, I is the environmental covariate Index, and the cc is the covariate coefficient to be estimated. The incorporation of covariates gives some flexibility from the constraint of the population stability assumption.

In addition, as this is intended as a multi-species modeling method, the software allows the user to implement predation using the Hollings' functional response types. Other types of functional responses have not yet been implemented.

# Uncertainties

The main uncertainties that affect the robustness of fisheries population models (and pose challenges to fisheries management) originate from a variety of sources including: observation error (associated with model input data and initial parameter estimates), dynamic model process errors, model structure misspecifications, and volatility in fishery socioeconomic dynamics. Few, if any models, can adequately deal with all of these uncertainties.

One of the main uncertainties with surplus production models lies within the simplicity of the population structure. The MSSPM has diagnostic capabilities to enable users to explore uncertainties by 1) examining parameter profiles, 2) performing retrospective analyses, and 3) forecasting with a range of uncertainties around parameter estimates.

# Peer reviews

The software was previously presented and the framework was reviewed to the WGSAM in 2021.

Models developed using the precursor to the MSSPM software, Kraken, were published including [Rob list MSPROD and other papers published here].

# Constructed Model

# Spatial/temporal resolution

Spatial resolution is discussed elsewhere with the Georges Bank data package.

Temporal resolution is an annual time step. The model time period is 1969-2019.

# Algorithm choices

For this constructed model of the Georges Bank Ecosystem, a range of population algorithms was used including 1) single species models, 2) single species models with covariates, 3) multi-species models with Type I predation, and 4) multi-species models with covariates. These models are discussed further in the Test scenarios section.

Within the MSSPM software, there are two sets of algorithms used for the parameter estimation:

| | |
|---|---|
| NLopt Library: | https://nlopt.readthedocs.io/en/latest/ |
| Bees Algorithm: | http://beesalgorithmsite.altervista.org/ |

Each is open source, available in C++ and freely available for download.

The NLopt Library is a library for nonlinear optimization. It contains support for both global and local algorithms as well as gradient-based and derivative-free algorithms. Each algorithm is described both in the software via online help and in the NLopt documentation pages here: https://nlopt.readthedocs.io/en/latest/NLopt_Algorithms/#comparing-algorithms.

The Bees Algorithm is a global search algorithm based on the behavior of honey bees. The “bees” explore the parameter space with more bees in various generations exploring further the parts of the parameter space that are more fruitful (i.e., have a better fitness value).

For the constructed model, the NLopt GN_ORIG_DIRECT_L minimizer algorithm was applied to the Maximum Likelihood objective function. The DIRECT (DIviding RECTangles) algorithms are deterministic search algorithms based on the splitting of the search domain into smaller and smaller hyper-rectangles.

# Assumptions

The major assumption of this model is that size/age-related dynamics are not necessary to adequately replicate the biomass time series for the given species. For models with environmental covariates, mechanisms of the environmental drivers’ effects are not modeled, only correlation evidence is provided. For models with predation, seasonal and spatial overlap are not considered. Process and observation error are not treated separately as would be the case for a state-space model. As this is a data-limited modeling approach being applied to a system that is not data limited, some data (e.g., age-structured catch) are ignored. The purpose of this model application is that it be used in a modeling suite or ensemble to allow exploration of uncertainty attributable to model structure.

# Data availability/software tools

The MSSPM is currently packaged with the Keyrun 2022 MySQL database. This can be found in the sample data distribution directory and should be imported into MSSPM once the user has started

MSSPM and created their project. MSSPM requires the MySQL database system be installed onto the user's machine. MSSPM has been ported to Linux and to Windows.

# Quality assurance/quality control

The software was previously presented and the framework was reviewed to the WGSAM in 2021.
For that review we implemented a process of 3-levels of simulation testing for this modeling software: comparison testing, integration testing, and skill assessment. Comparison testing used simulated data from an operating model with the same structure as used in the software with known parameter values, known harvest time series, and known biomass time series (without and with uncertainty). Integration testing uses simulated data from an operating model with a different structure (Norwegian Barents Atlantis Model, NOBA) thus different parameters, and known catch and biomass time series without uncertainty to demonstrate that software modules (e.g., parameter estimation, diagnostics, forecasting/projection) produce reasonable results. The skill assessment uses simulated data from an operating model with a different structure (NOBA) and different parameters to create catch and biomass index time series with uncertainty. For comparison testing, MSSPM fits the operating model perfectly ($r^2$=1.00) for data without uncertainty, and fits decrease as uncertainty increases in biomass time series (e.g., with cv=0.5, $r^2$ ranges from 0.50 to 0.60). For integration testing, the best fit model $r^2$ was 0.44. Skill assessment model runs are ongoing and have not been analyzed at this date.

With updates to the software, additional comparison testing has been conducted, i.e., the software has been tested with simulation data. That is, based upon the current model desired, the software can generate either relative or observed biomass. This biomass is then run through the model, and the results should be (and are) a perfect fit. Random error may also be added to the simulation data to check the model's behavior under uncertainty.

# Test scenarios

All models are developed using catch data to drive the dynamics and survey indices (relative biomass) data for fitting (). The time period of the model is 1969-2019. The species modeled and their trophic guilds are listed in Table 1.

*Figure 2. Catch and relative biomass time series used for the Georges Bank MSSPM.*

| Species Name | Guild |
|---|---|

| | |
|---|---|
| Atlantic cod | Groundfish |
| Atlantic herring | Small pelagics |
| Atlantic mackerel | Small pelagics |
| Goosefish | Groundfish |
| Haddock | Groundfish |
| Silver hake | Groundfish |
| Spiny dogfish | Elasmobranchs |
| Winter flounder | Flatfish |
| Winter skate | Elasmobranchs |
| Yellowtail flounder | Flatfish |

*Table 1. Georges Bank Fisheries species models using MSSPM.*

Four test scenarios were conducted to compare basic surplus production models with and without environmental covariates and predation. The model names, descriptions, and parameters estimated are given in Table 2.

| Model Name | Description | Parameters estimated |
|---|---|---|
| LogisticFullEst | Base model - with Logistic growth, predation, and environmental covariates | B0,r,K, cc, ρ |
| LogisticSS | Multiple single species models - with Logistic growth | B0,r,K |
| LogisticCov | Multiple single species models with - Logistic growth environmental covariates | B0,r,K, cc |
| LogisticPred | Base model - with Logistic growth and predation | B0,r,K, ρ |

*Table 2. Candidate models for Georges Bank Ecosystem using MSSPM.*

The model equation for the full model is

$$B_{i,t+1} = B_{i,t} + r_i B_{i,t} \left(1 - \frac{B_{i,t}}{K_i}\right) - C_{i,t} - B_{i,t} \sum \quad \rho_{i,j} B_{j,t} \qquad \text{Equation 2}$$

where, B is biomass (in metric tons), r is population growth rate, K is carrying capacity, C is catch (in metric tons), ρ is a predation coefficient, i is focal species index, j is predator species index, and t is time (year). For models with environmental covariates, Equation 1 is applied to r or K. B0, r, and K are estimated for all species. Environmental covariate and predation coefficients are only applied to select species, which are identified in the parameter estimates for the most parsimonious model given in Appendix C.

Models listed in Table 2 were fit to relative Biomass indices using the Maximum Likelihood objective function and the GN_DIRECT_L minimizer algorithm. Akaike Information Criterion was used to select the most parsimonious model from the candidate set.

# Corroboration with observation

The most parsimonious model (i.e., model with the lowest AIC value) included predation and environmental covariates applied to some species (Table 3). A summary of fit statistics for this model is given in Appendix D.

| ModelName | $r^2$ | SSResiduals | AIC |
|---|---|---|---|
| LogisticFullEst | 0.244 | 3.39E+16 | 1193.725 |
| LogisticPred | 0.227 | 3.40E+16 | 1198.364 |
| LogisticSS | 0.285 | 7.19E+16 | 1201.359 |
| LogisticCov | 0.326 | 1.34E+17 | 1202.010 |

*Table 3.Model selection for Georges Bank Ecosystem using MSSPM.*

The model was able to reasonably replicate historical biomass trends for each species, guild, and the system (Figure 2-4).

*Figure 3. Model fit to species biomass.*

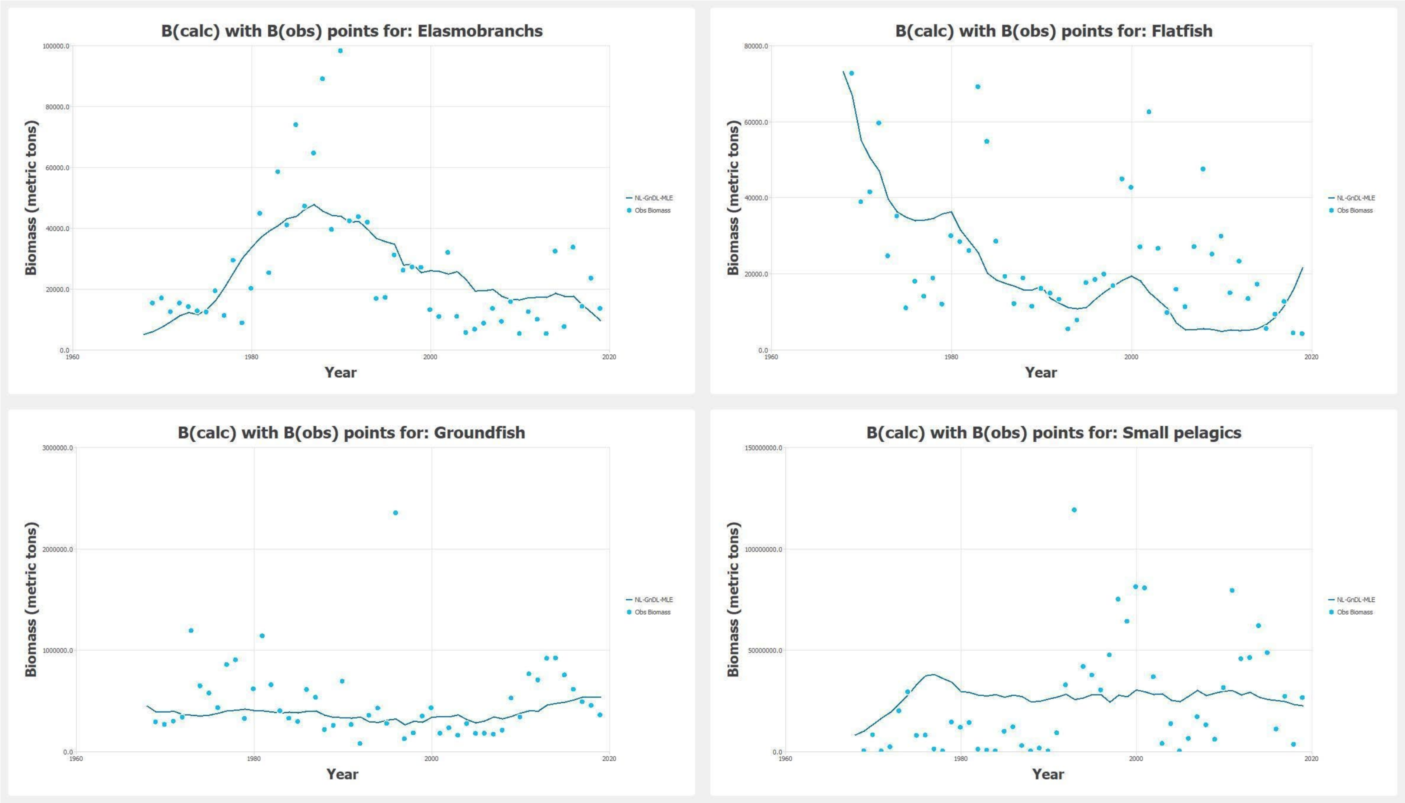


*Figure 4. Model fit to guild biomass.*

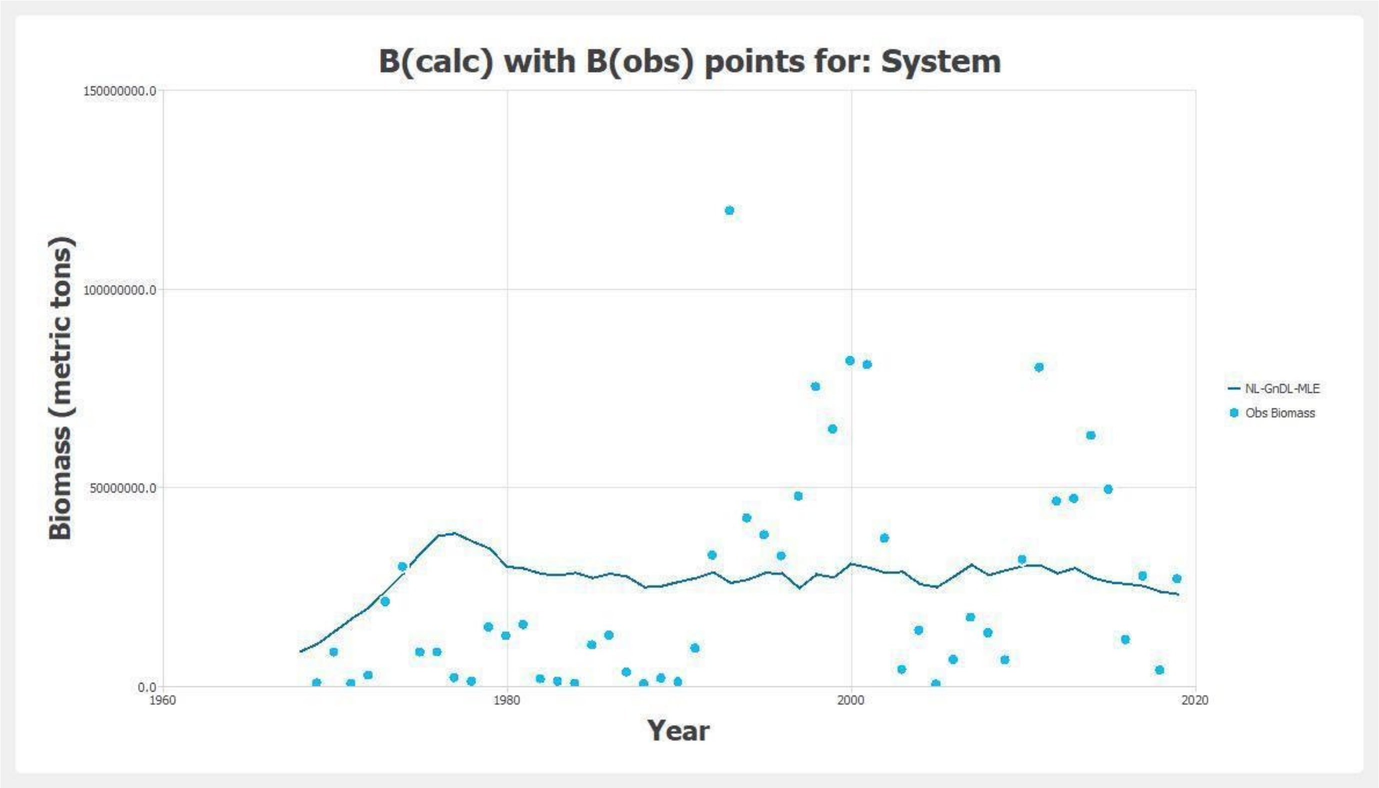


*Figure 5. Model fit to system biomass.*

# Uncertainty/Sensitivity

To explore the uncertainty associated with the parameters and model structure, parameter fitness profiles were created and retrospective analysis was conducted. As parameter space is multidimensional, single dimension profiles may have limited use. More information on how these analyses were produced are given in Appendix A – Model Diagnostics section.

## Parameter Profiles

Parameter profiles were created by calculating fit for 50 different parameter values at ±50% of the estimated value. Two-dimensional plots for single parameters are provided in Figures 5-7. Three-dimensional plots of parameter combinations are created in the software for this model.

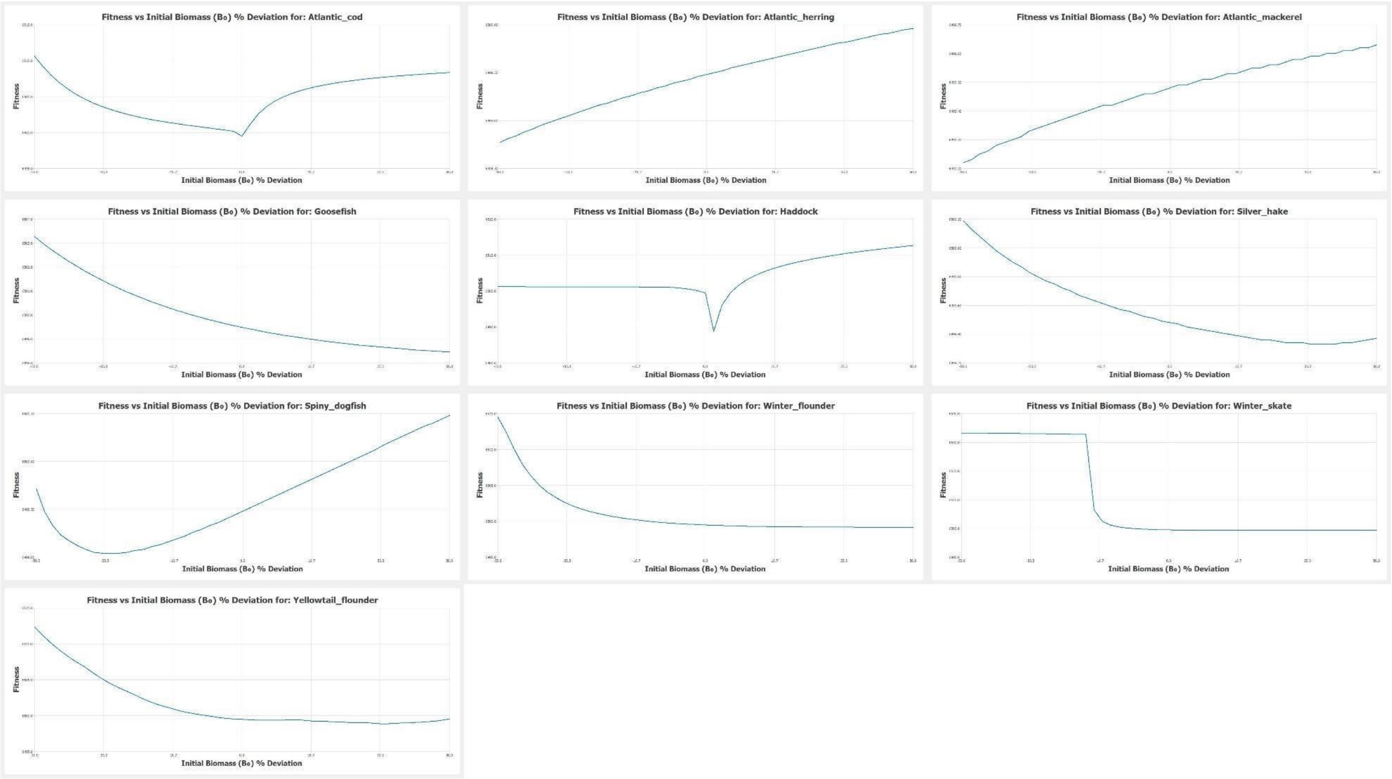


*Figure 6. $B_0$ profiles*

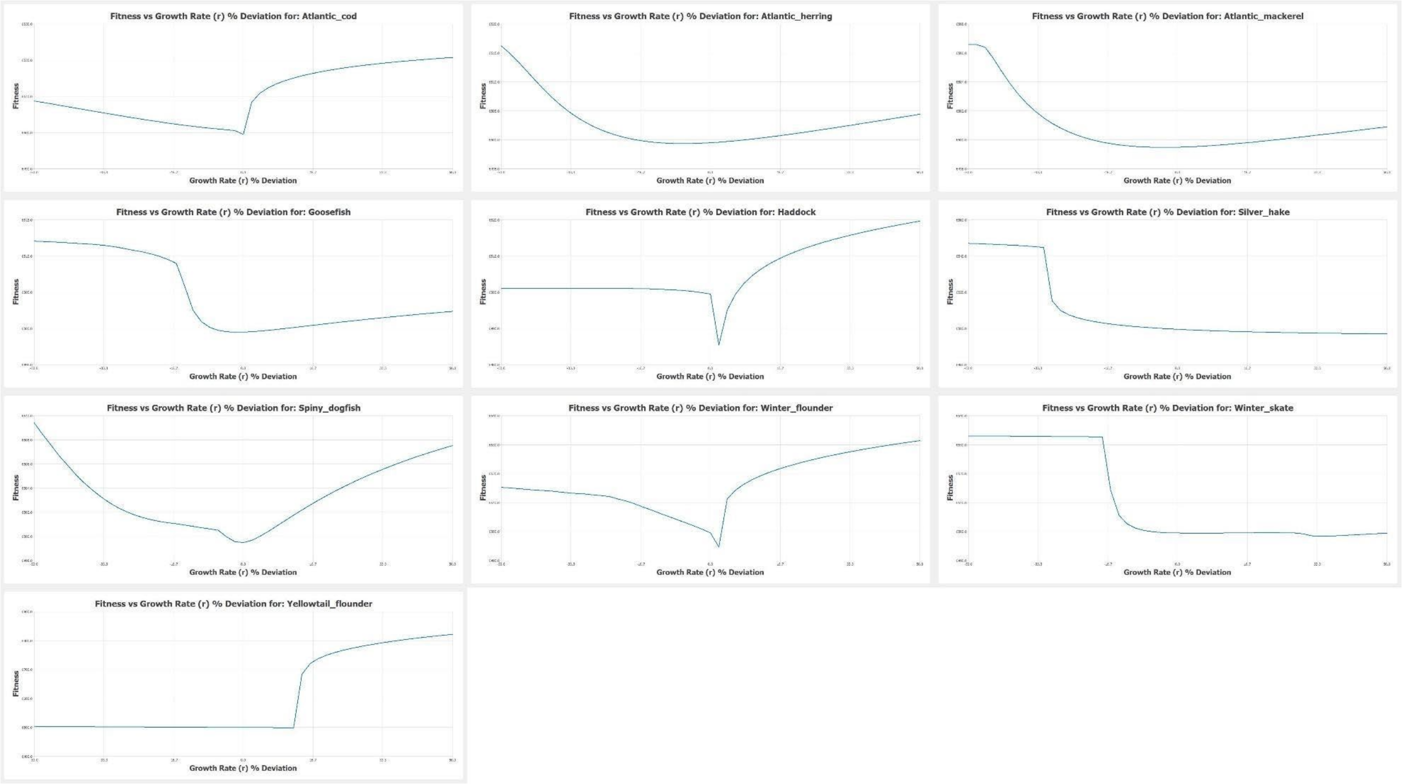


*Figure 6. r profiles*

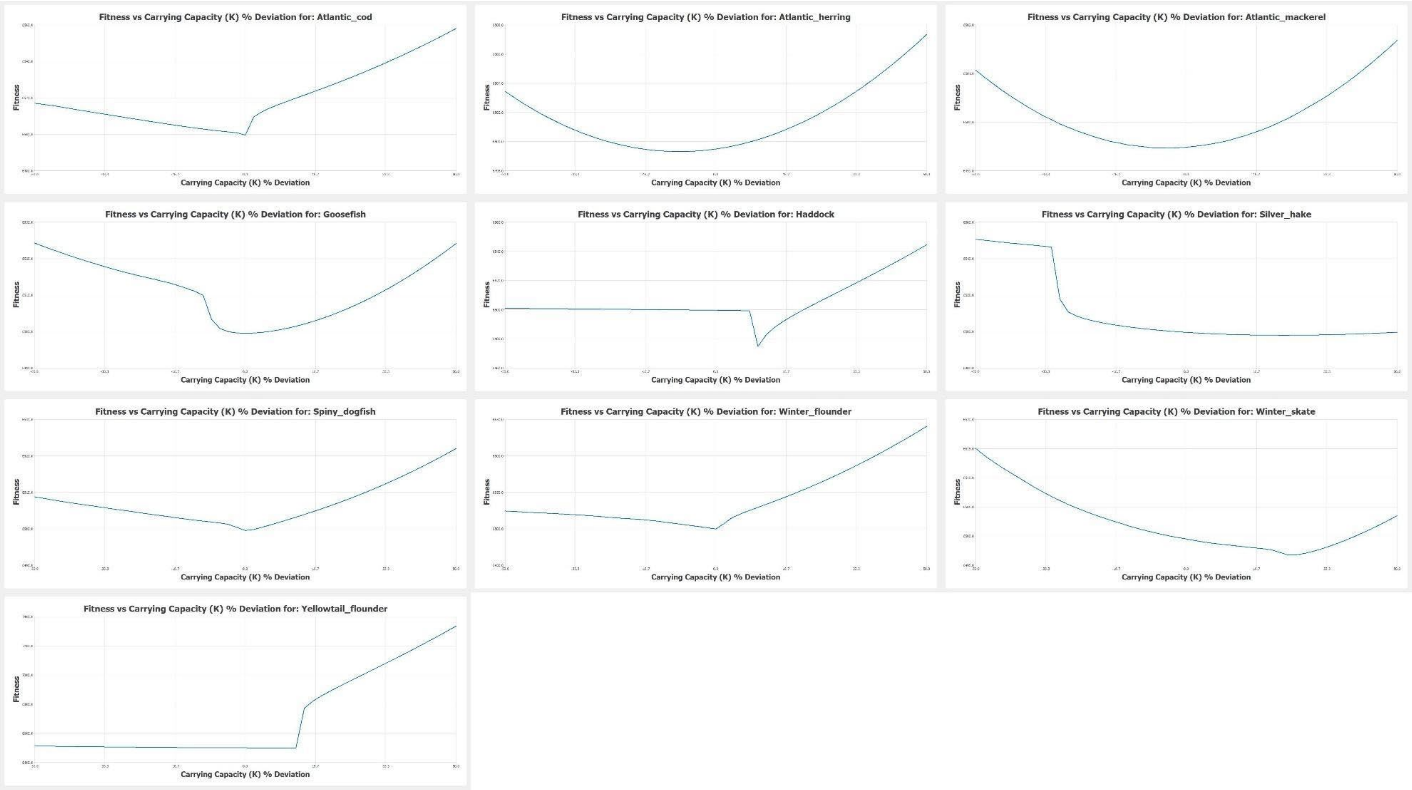


*Figure 7. K profiles*

# Retrospective Analysis

Retrospective analysis plots were created by four one-year peels. Mohn's rho values are provided in Table 4.

*Figure 8. Retrospective analysis*

|  | Atlantic cod | Atlantic herring | Atlantic mackerel | Goosefish | Haddock | Silver hake | Spiny dogfish | Winter flounder | Winter skate | Yellowtail flounder |
|---|---|---|---|---|---|---|---|---|---|---|
| Mohn's Rho (Initial Absolute Biomass) | 0.002 | 0.00E+00 | 0.00E+00 | 0 | 0 | 0.00E+00 | 0 | -0.001 | 0 | 0 |
| Mohn's Rho (SurveyQ) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mohn's Rho (Growth Rate) | 0 | 0.014 | -0.011 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mohn's Rho (Carrying Capacity) | 0 | 0.00E+00 | -3.30E-02 | -0.01 | 0 | 0.00E+00 | -0.001 | 0 | 0 | 0 |
| Mohn's Rho (Predation Rho) | 0 |  |  | 0 |  | 0.004 | 0 |  | 0.064 |  |
| Mohn's Rho (Estimated Biomass) | 0.02 | 2.00E-02 | 2.00E-02 | 0.02 | 0.02 | 2.00E-02 | 0.02 | 0.02 | 0.02 | 0.02 |

*Table 4. Mohn’s rho values for keyrun species.*

# Peer review

The software was previously presented and the framework was reviewed to the WGSAM in 2021.

Models developed using the precursor to the MSSPM software, Kraken, were published including [Rob list MSPROD and other papers published here].

# Appendix A. MSSPM Software Overview

The MSSPM software is designed to allow fishery analysts to estimate aquatic ecosystem population parameters. These parameters may then be used to forecast future biomass trends. Care has been taken to design an intuitive and functional user experience. The forecast component, REMORA, is designed with a select group of user interface widgets that enable a stakeholder to quickly interact with the estimated model parameters and to construct a forecast using easy-to-tweak forecast parameters.

## Layout

The user interface consists of movable panels allowing the user to customize their layout. The default layout allows for program flow from left to right, with a navigator panel being on the left, data input and model control in the middle, and output results on the right. The panel at the bottom displays real-time feedback for the user as the model runs as well as making a session log file visible.

## Philosophy & General Principles

The MSSPM software is built using open-source and cross-platform libraries and systems. It uses Qt's library of graphical user interface (GUI) widgets as its basic framework, taking advantage of Qt's messaging capabilities (i.e., signals/slots) for inter-GUI communication. It uses MySQL as its relational database management system with 85 tables comprising the data storage functionality. The user saves data to the database as they move from one input data tab to the next. The user may also import/export data from/to data files on disk.

In order to balance numeric precision with ease-of-reading, MSSPM uses a scheme of significant digits. This optional feature controls how many significant digits are defined for the data in the GUIs. This greatly increases the readability of the data tables.

MSSPM's logging functionality captures various software checkpoints, warnings, and errors and writes them to a timestamped data file.  The user has the capability to view the current session's log file or that of a previous session.  Log files are periodically deleted upon user approval.

MSSPM's inline help functionality is based upon Qt's tooltips and What's This? help.  The former displays to the user a brief help message upon hovering over a graphical widget. The latter is enabled by the user by first clicking on the What's This? help icon in the toolbar and then clicking on a widget for more information. If available, more detailed help will appear. Finally, under the Help main menu item, the user can find the About item. This item describes the current MSSPM version in detail, including the run time, version number, build data, and version numbers of all 3rd party software used.

# Functionality

The MSSPM allows the user to create a population model, estimate model parameters, run parameter diagnostics, and generate biomass forecasts. The user chooses from a library of 3rd party, open source optimization algorithms for the parameter estimation.

# Creating a Population Model

The MSSPM allows the user to create a custom population numerical model, which defines species biomass values, based upon 4 terms: Growth Term, Harvest Term, Competition Term, and Predation Term. The first term is additive whereas the other terms are subtractive. In addition to the user creating their own model, they may select from a list of preset models (i.e., Schafer Model).

After the user defines their model, the appropriate data input GUIs will be enabled and the user proceeds to input their data. Care should be taken by the user to enter data using consistent units. For example, biomass units should be in metric tons and be consistent between all species.

The Growth term defines how the population grows. It’s a function of growth rate of species i, r(i) (a unit-less term), the calculated biomass of the species, B(i) (metric tons) at time t, and possibly a carrying capacity (maximum biomass of species which can be supported by the environment) K(i) term (also in metric tons). Available choices are:

- Linear: $r(i)B(i,t)$
- Logistic: $r(i)B(i,t)(1-B(i,t)/K(i))$

The Harvest term defines how the population is decreased by fishing. Available choices are:

- Effort: $q(i)E(i,t)B(i,t)$
- Catch: $C(i,t)$

The estimated parameter q(i) is the unit-less catchability parameter. The non-estimated input parameter E(i,t) is the unit-less effort data. Both parameters are typically between 0.0 and 1.0.

The Competition term describes how biomass is decreased by fish competing against other fish for the same food. Available choices are:

- NO_K: $B(i,t)\sum\alpha(i,j)B(j,t)$
- MS-PROD: $r(i)B(i,t)[(\sum\beta(i,j)B(j,t))/K(G) - (\sum\beta(i,G)B(G,t))/(K(\sigma) - K(G))]$

The estimated parameters are: $\alpha(i,j)$ - the effect of species j on species i, $\beta(i,j)$ - the effect of species j in same guild as species i, on species i, $\beta(i,G)$ - the effect of guild G on species i, $K(G)$ - guild carrying capacity, and $K(\sigma)$ - the system carrying capacity. A guild is defined as a group of species.

The Predation term describes how the biomass is decreased by fish eating other fish. Available choices are:

- Type I: $B(i,t)\sum\rho(i,j)B(j,t)$
- Type II: $B(i,t)\sum[\rho(i,j)B(j,t)/(1+\sum h(k,j)\rho(k,j)B(k,t))]$
- Type III: $B(i,t)^{b_i+1}\sum[\rho(i,j)B(j,t)/ (1+\sum h(k,j)\rho(k,j)B(k,t)^{b_k+1})]$

The estimated parameters are: $\rho(i,j)$ - the effect of predator species j on prey species i, $h(k,j)$ - handling time for predator species j with prey species k, and $b(k)$ - a predator dependent parameter.

An example model equation for logistic growth, catch harvest, and NO_K competition is:

$$B(i,t+1) = B(i,t) + r(i)B(i,t)(1 - B(i,t)/K(i)) - C(i,t) - B(i,t)\sum\alpha(i,j)B(j,t)$$

Note that this model generates a time series of biomass for species i and j over time t. The time series requires an initial biomass (given or estimated) at t=0.

# Estimating Model Parameters

After the user has entered all population model data, including min/max ranges for each parameter to be estimated, they then set the model algorithm parameters that will determine how the model is to run. An estimation model is run until 1 of 3 stop conditions are met: the fitness value reaches a certain value, after time t, or after a specific number of function evaluations. The fitness value measures how well the estimated biomass fits the observed biomass as described below.

The model algorithm parameters are: objective criterion, estimation algorithm, minimizer algorithm, scaling algorithm, and stopping conditions. Each has a variety of settings which may yield different results. The user should try several distinct combinations as they estimate their parameters. Users may create an "ensemble" of runs with each run consisting of sub runs with unique model algorithm parameter settings. The user may then average the results.

The Objective Criterion parameter measures how well (i.e., fitness) the estimated biomass fits the observed biomass or relative biomass. Available choices are: least squares, maximum likelihood, or model efficiency.

The Estimation Algorithm is currently defined as being one of two families of algorithms: the 3rd party, open source NLopt non-linear optimization library or the 3rd party, open source Bees Algorithm, modeled after a global search algorithm derived from the behavior of honey bees.

The Minimizer Algorithm is the sub algorithm selected from the family of Estimation Algorithms selected. There is only one Bees Algorithm available but there are several (global/local and deterministic/stochastic)

NLopt optimization algorithms available. Each algorithm is defined in more detail from the inline What's This? help. When running a stochastic algorithm, each time that the estimation is run, a random seed value will be used. If the user wishes to run stochastically, but also wishes to reproduce the results, a box may be checked to set the initial seed to a constant. In this fashion, the stochastic runs will use the same random seeds each time they're run. More details on specific optimizing algorithms are provided in the MSSPM User Guide.

The Scaling Algorithm is used to rescale the biomass data since biomass values for each species may not be scaled the same. The user may select from a list of scaling algorithms. Available choices are:

- mean: $(B - B_{ave}) / (B_{max} - B_{min})$
- min/max: $(B - B_{min}) / (B_{max} - B_{min})$

The Stopping Parameters are used to terminate the estimation evaluations. There are 3 stopping parameters available and 1 or more may be used. The estimation will stop when one of the following conditions is met: after reaching a specific value, after time t, or after a specific number of function evaluations. The user may "watch" the estimation's fitness values change over time from the Progress window and adjust the stopping parameter(s) accordingly.

The user may create an "ensemble" of runs consisting of multiple runs using the same or different algorithms. An ensemble's run sub runs are then averaged (either by parameter or by biomass) and the averaged estimated biomass is shown against the individual runs.

The user can add to a "review list" the results from a run to view with other run results in a tabular format and generate a report from a selected run. In this fashion the user can compare previous runs and also "call-up" a previous run as each run's initial parameter settings are saved.

# Model Diagnostics

The user may visually inspect the estimated parameter's neighborhood using a 2d chart (one parameter at a time) or using a 3d surface plot (2 parameters at a time). By varying the number of points and percent variation on either side of the estimated parameter, the user can visualize the parameter space in question and see if the estimated parameter is a global minimum point for the parameter(s) in question.

Additionally, a retrospective analysis (i.e., Mohn's Rho) may be run with the user selecting the number of years to be "peeled" from subsequent runs. The Mohn's Rho is a numerical measurement of the severity of a retrospective pattern [Deroba 2014]. Its value is given by:

- Mon's Rho: $\{\Sigma \left[ (X(t-n,t-n) - X(t-n,t)) / X(t-n,t) \right] \} / x$

where Σ goes from n=1 to x years. The value n represents the number of years "peeled off" from the complete time series.

# Running a Forecast

With estimated parameters found, the user may now project forward the population model to see how the population may change in the future. The data needed for a forecast are: the number of years/run, number of runs/forecast, the harvest data to be used for the duration of the forecast, and the % error desired around each parameter. Multiple forecasts are run in a Monte Carlo fashion using the preceding data.

The forecast display consists of all of the individual runs in gray (i.e., those using % errors in their parameter values) and one overlaid run in blue with 0% error. In this fashion, the user can quickly see how the Monte Carlo runs compare with the 0% error run.

# Running REMORA

REMORA is a streamlined forecasting interface for stakeholders (see next section for more details) that allows for quick forecast iterations. It is designed to be used after the user has completed their estimation and is ready to run different forecast scenarios.

# Analyzing Results

The following output charts are available in the Output Panel: Biomass vs Time, Harvest vs Time, and Exploitation Rate. Data may be viewed per species, per guild, or per system. Additionally, maximum sustained yield lines may be superimposed over the output plots.

# Generate Simulated Data

For testing purposes, the software can generate its own simulated biomass data, with user inputted percent error if desired. The simulated biomass data are generated with the current parameter settings. In this fashion, the user can confirm that the estimation is working correctly by confirming, for example, that with simulated biomass data with 0% error, a perfect fit of the estimated parameters is found.

## REMORA

The REMORA (REsource Management Option Review and Analysis) interface is designed for the interactive forecasting use of a stakeholder after parameter estimations have been completed. The user interface is a slimmed down version of the main application's forecast controls. It gives the stakeholder intuitive control for the tweaking of uncertainty parameters as subsequent forecasts are created. For example, in lieu of harvest values having to be entered in tabular format, an interactive plot has been developed where the stakeholder can graphically construct the desired harvest trend throughout the forecast period.

The REMORA workflow is to generate a forecast and then allow users to interactively adjust errors around parameters and harvest values and create another forecast. Through this iterative process the user can then define harvest values and error values that yield a desired forecast, thus a suitable policy to consider.

# Components

The MSSPM is written in C++ and built with the following open-source tools. The Qt[1] graphical user interface widget toolkit is used as the GUI framework. The MySQL[2] relational database system is used to store all model data. Model estimation is performed by the NLopt[3] nonlinear optimization library. In addition to the algorithms provided by NLopt, an implementation of a Bees algorithm[4], a global search algorithm based upon honey bee behavior, has been provided for the analyst as well. Linear algebra operations are performed with the Boost[5] library. Finally, two other modules were used: QDarkStyleSheet[6] for dark widget colors when in dark mode, and linuxdeployqt[7] to facilitate creating a Linux release of MSSPM.

The specific version number of each of the aforementioned libraries, as well as a link to their websites, may be found in the About dialog from MSSPM's Help menu group.

---

[1] https://www.qt.io

[2] https://www.mysql.com

[3] https://nlopt.readthedocs.io

[4] http://beesalgorithmsite.altervista.org

[5] https://www.boost.org

[6] https://github.com/ColinDuquesnoy/QDarkStyleSheet

[7] https://github.com/probonopd/linuxdeployqt

# Appendix B. Bottom Temperature Time series

Author and Analyst Hubert duPontavice

## Data Sources

Three bottom temperature products to get temperature from 1959 to 2019:

1. Numerical simulation of the NWA Ocean was performed with the Regional Ocean Modelling System (ROMS)

   PERIOD: 1958-1992

   INITIAL RESOLUTION: ~7km → REGRID: 1/10°

   (Shchepetkin and McWilliams, 2005)

   https://www.sciencedirect.com/science/article/pii/S1463500304000484

2. Global Ocean Physics Reanalysis (Glorys reanalysis)

   PERIOD: 1993-2019

   INITIAL RESOLUTION: 1/12° → REGRID: 1/10°

   (Fernandez and Lellouche, 2018; Lellouche et al., 2018)

   https://resources.marine.copernicus.eu/product-detail/GLOBAL_REANALYSIS_PHY_001_030/INFORMATION

3. Furthermore, we use bottom temperature from the Northwest Atlantic Regional Climatology to estimate a monthly decadal bias (Seidov et al., 2016a, 2016b)

INITIAL RESOLUTION: 1/10°

https://www.ncei.noaa.gov/products/northwest-atlantic-regional-climatology

# Data Processing

Bottom temperature data from Glorys reanalysis was not processed.

Bottom temperatures from ROMS-NWA (used for the period 1958-1992) were bias-corrected.

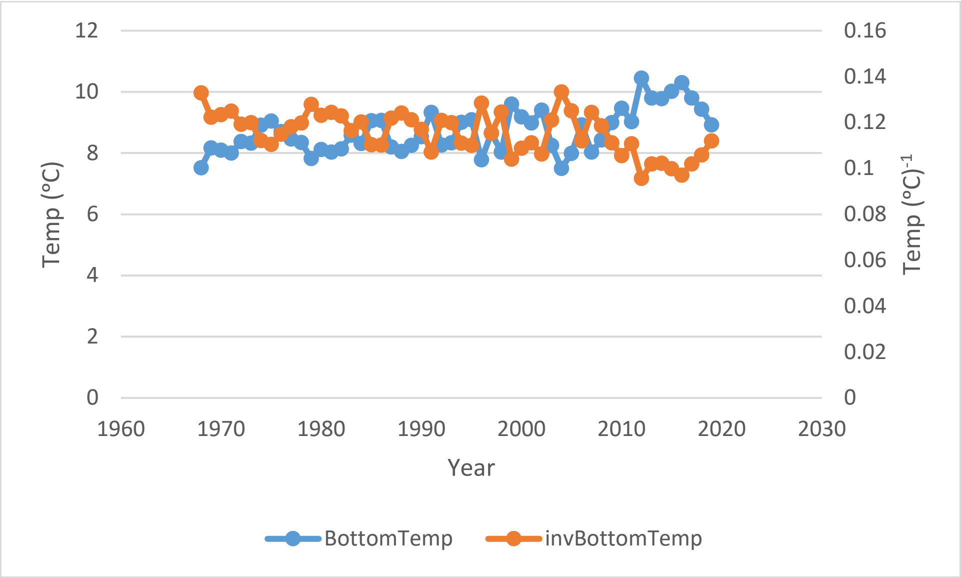
Previous studies that focused on the ROMS-NWA-based Cold Pool highlighted strong and consistent warm bias in bottom temperature of about 1.5˚C during the stratified seasons over the period of 1958-2007 (Chen et al., 2018; Chen and Curchitser, 2020) In order to bias-correct bottom temperature from ROMS-NWA, we used the monthly climatologies of observed bottom temperature from the Northwest Atlantic Ocean regional climatology (NWARC) over decadal periods from 1955 to 1994. The NWARC provides high resolution (1/10° grids) of quality-controlled in situ ocean temperature based on a large volume of observed temperature data (Seidov et al., 2016a, 2016b)

(https://www.ncei.noaa.gov/products/northwest-atlantic-regional-climatology). The first step was to re-grid the ROMS-NWA to obtain bottom temperature over the same 1/10° grid as the NWARC. Then, a monthly bias was calculated in each grid cell and for each decade (1955–1964, 1965–1974, 1975–1984, 1985–1994) in the MAB and in the SNE shelf:

Where T(i,d) Climatology is the NWARC bottom temperature in the grid cell i for the decade d and T(i,d) ROMS-NWA is the average ROMS-NWA bottom temperature over the decade d in the grid cell i.

# Bottom Temperature Index

A GIS Shapefile for Georges Bank was used to extract relevant grid cells for the model. Cell values were averaged by year to produce an annual index. The inverse of the index was calculated for application to species where growth rate or carrying capacity might be inversely related to the bottom temperature.



# Bibliography

Chen, Z., Curchitser, E., Chant, R., and Kang, D. 2018. Seasonal Variability of the Cold Pool Over the Mid-Atlantic Bight Continental Shelf. Journal of Geophysical Research: Oceans, 123: 8203–8226.

Chen, Z., and Curchitser, E. N. 2020. Interannual Variability of the Mid-Atlantic Bight Cold Pool. Journal of Geophysical Research: Oceans, 125. https://onlinelibrary.wiley.com/doi/10.1029/2020JC016445 (Accessed 13 January 2021).

Fernandez, E., and Lellouche, J. M. 2018. Product user manual for the global ocean physical reanalysis product GLORYS12V1. Copernicus Product User Manual, 4: 1–15.

Gamble, R.J., and Link, J.S. 2009. Analyzing the tradeoffs among ecological and fishing effects on an example fish community: a multispecies (fisheries) production model. Ecol. Model. 220: 2570−2582

Gamble, R.J., and Link, J.S. 2012. Using an aggregate production simulation model with ecological interactions to explore effects of fishing and climate on a fish community. Mar Ecol Prog Ser. 459: 259-274

Lellouche, J.-M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., Benkiran, M., et al. 2018. Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1⁄12° high-resolution system. Ocean Science, 14: 1093–1126.

Seidov, D., Baranova, O. K., Johnson, D. R., Boyer, T. P., Mishonov, A. V., and Parsons, A. R. 2016a. Northwest Atlantic Regional Climatology (NCEI Accession 0155889). NOAA National Centers for Environmental Information. https://www.ncei.noaa.gov/archive/accession/0155889 (Accessed 25 March 2021).

Seidov, D., Baranova, O. K., Boyer, T., Cross, S. L., Mishonov, A. V., and Parsons, A. R. 2016b. Northwest Atlantic regional ocean climatology.: 3.2 MB. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, National Centers for Environmental Information.

Shchepetkin, A. F., and McWilliams, J. C. 2005. The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. Ocean Modelling, 9: 347–404.

# Appendix C. Final model parameter estimates from most parsimonious model

## $B_0$, r, and K

| Species | $B_0$ | r | K |
|---|---|---|---|
| Atlantic cod | 182500 | 0.297 | 206000 |
| Atlantic herring | 7591000 | 0.900 | 54340000 |
| Atlantic mackerel | 575000 | 0.978 | 3615000 |
| Goosefish | 41200 | 0.291 | 40640 |
| Haddock | 25830 | 0.357 | 146700 |
| Silver hake | 198300 | 0.900 | 244900 |
| Spiny dogfish | 2918 | 0.291 | 20530 |
| Winter flounder | 31880 | 0.510 | 31480 |
| Winter skate | 2128 | 0.553 | 20850 |
| Yellowtail flounder | 41240 | 0.461 | 50780 |

## cc

| Species | K | r | K· | r· |
|---|---|---|---|---|
| Atlantic cod | 0.0 | 0.0 | | |
| Atlantic herring | 0.911 | 0.0 | Bottom temp | |

| | | | | |
|---|---|---|---|---|
| Atlantic mackerel | 1.198 | 0.0 | Bottom temp | |
| Goosefish | 0.0 | 0.0 | | |
| Haddock | 0.699 | 0.0 | Bottom temp | |
| Silver hake | 0.927 | 0.0 | Bottom temp | |
| Spiny dogfish | 0.0 | 0.00330 | | Bottom temp$^{-1}$ |
| Winter flounder | 0.0 | 0.0333 | | Bottom temp$^{-1}$ |
| Winter skate | 0.791 | 0.0 | Bottom temp | |
| Yellowtail flounder | 0.167 | 0.0 | Bottom temp$^{-1}$ | |

$\rho$

| | **Prey** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Predator** | Atlantic cod | Atlantic herring | Atlantic mackerel | Goosefish | Haddock | Silver hake | Spiny dogfish | Winter flounder | Winter skate | Yellowtail flounder |
| Atlantic cod | 0 | 1.77E-07 | 6.02E-08 | 0 | 0 | 1.31E-08 | 0 | 0 | 0 | 0 |
| Atlantic herring | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Atlantic mackerel | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Goosefish | 0 | 1.45E-07 | 4.92E-07 | 0 | 0 | 8.99E-08 | 0 | 0 | 0 | 0 |
| Haddock | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Silver hake | 0 | 2.36E-06 | 2.67E-06 | 0 | 0 | 9.53E-07 | 0 | 0 | 0 | 0 |
| Spiny dogfish | 0 | 6.44E-07 | 4.90E-07 | 0 | 0 | 3.25E-08 | 0 | 0 | 0 | 0 |
| Winter flounder | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Winter skate | 0 | 1.44E-07 | 1.52E-07 | 0 | 0 | 5.20E-08 | 0 | 0 | 0 | 0 |
| Yellowtail flounder | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

# Appendix D. Model Fit Summary

| Statistic | Atlantic cod | Atlantic herring | Atlantic mackerel | Goosefish | Haddock | Silver hake | Spiny dogfish | Winter flounder | Winter skate | Yellowtail flounder | Model |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SSresiduals | 5.05E+11 | 3.36E+16 | 3.48E+14 | 3.89E+10 | 8.16E+10 | 6.05E+12 | 6.56E+09 | 6.53E+09 | 7.62E+09 | 4.18E+09 | 3.39E+16 |
| SSdeviations | 2.05E+11 | 1.57E+15 | 9.55E+12 | 4.21E+09 | 1.14E+11 | 3.84E+11 | 1.80E+09 | 3.21E+09 | 3.09E+09 | 4.29E+09 | 1.58E+15 |
| SStotals | 7.09E+11 | 3.52E+16 | 3.57E+14 | 4.31E+10 | 1.95E+11 | 6.44E+12 | 8.36E+09 | 9.75E+09 | 1.07E+10 | 8.47E+09 | 3.55E+16 |
| $r^2$ | 0.288 | 0.045 | 0.027 | 0.098 | 0.582 | 0.06 | 0.215 | 0.33 | 0.288 | 0.506 | 0.244 |
| r | 0.562 | 0.174 | 0.024 | 0.462 | 0.756 | 0.213 | 0.459 | 0.312 | 0.515 | 0.514 | 0.399 |
| AIC | 1203.794 | 1781.27 | 1543.593 | 1070.531 | 1109.023 | 1332.979 | 977.989 | 977.729 | 985.767 | 954.569 | 1193.725 |

# Annex 4: Hydra Model

## Hydra Methology

### Introduction

The original purpose of the Hydra simulation model was to create simulated data for testing simpler multispecies assessment models (Gaichas et al. 2017). Further model development repurposed Hydra as an MSE tool to test proposed multispecies and ecosystem level management procedures to support EBFM analyses for the New England Fishery Management Council, and were reviewed in 2018 as part of an Ecosystem Based Management Strategy Review for Georges Bank.

The current Hydra estimation model can still be used for the above purposes (no simulation code was removed), but with the potential added benefit of estimating operating model parameters by fitting to data from the ecosystem, rather than iterative hand calibration.

Multispecies models "of intermediate complexity" between single population and full ecosystem models potentially combine the best aspects of current single species assessment models (including demographic structure and statistical parameter estimation) with key ecological linkages between species (Plagányi et al. 2012, Collie et al. 2014). The Northeast Fisheries Science Center (NEFSC) is currently applying multispecies production and delay-difference models to a simulated system coarsely modeled after the Georges Bank fish community. The focus of this work is to evaluate the performance of a length-based multispecies approach for providing management advice. While we do not yet have a fitted model ready to provide catch advice, this review should identify next steps for further developing model code, inputs, and diagnostics towards this goal.

To evaluate Hydra's performance as an estimation model, we converted it from a simulation model by incorporating likelihood functions to fit the model to time series of aggregate fishery catch, survey biomass, size compositions from the fishery and survey, and length-specific predator diets. Assessment model performance evaluation is in progress by comparing "true" biomass, abundance, and fishing mortality rates to model-estimated quantities using skill assessment metrics (e.g., correlation, RMSE, Modeling efficiency; Stow et al. 2009). Finally, the length-based multispecies model is being fit to actual observations from Georges Bank to evaluate operational potential within a particular region.

### Methodology

The reviewed model is based on the simulation model Hydra (Gaichas et al. 2017), implemented in ADMB (Fournier et al. 2012). Hydra's structure is derived from the length-based multispecies simulation model, LeMANS (Hall et al. 2006; Rochet et al. 2011), with additional options for growth, and recruitment functional forms and more detailed fishing fleets.

A key feature of Hydra that differentiates it from LeMANS and other length-based models is its treatment of length information. Attempting to balance model complexity and run time with available data and adequately modeling length-based processes, Hydra applies an equal number of length bins to all modeled species. The modeler specifies the overall number of bins (5 in the original simulation model), and the size in cm spanned by each bin for each species. The original Hydra simulation model used narrower length bins for smaller fish sizes and a length bin spanning a wider range of sizes for larger fish to efficiently represent changes size selectivity of predators and fisheries. The estimation model uses equal width bins equally dividing each species' maximum observed length into the overall number of bins.

Modeler decisions on length bin number and width combined with individual species growth parameters both contribute to the model temporal scale, which internally calculated and set equal to the time required for the fastest-growing species to grow out of its smallest length bin (if that is less than one year). Therefore, model timesteps will never exceed one year, but can be subannual.

For this review, the simulation model code was forked and an objective function added, with modifications to the data section to accommodate new data inputs and to the parameter section to estimate fishery and survey selectivity and catchability. As an estimation model, Hydra can be fitted to multiple data streams. Currently the objective function comprises the following components, catch (by species for each fleet), size composition of the catch (by species for each fleet), survey abundance index (by species for each survey), survey size composition (by species for each survey), survey diet composition (stomach weight by predator size bin for each predator species and survey), and recruitment (annual deviations from mean recruitment by species). A detailed description of the objective function equations is available here.

Hydra input files were developed directly from `mskeyrun` datasets by modifying functions in the `hydradata` R package. The function `create_Rdata_mskeyrun.R` (code) allows the user to specify whether datasets should be constructed from Atlantis-simulated or real Georges Bank data, and the number of length bins to use for composition data, then creates an R data object. This data object is then used to create data and parameter input files using the function `hydradata::create_datpin_files()`.

Hydra output visualizations were developed to evaluate model fits with different settings and input datasets. The current diagnostics include model fits to each survey index, model fits to catch by fleet, comparisons of estimated and observed length compositions by year and in aggregate across years for each modeled species, residuals of fits to length compositions, and comparisons of estimated and observed diet composition. In addition, model outputs of estimated recruitment, total biomass, fishery and survey selectivity, fishing mortality, and predation mortality are produced. Initial model skill assessment compares model output total biomass with known total biomass for the simulated datasets.

All model code, data processing, and visualization is available online:

- Hydra model estimation version code https://github.com/thefaylab/hydra_sim
- Hydra estimation model data file generation https://github.com/thefaylab/hydradata
- Hydra model diagnostics https://github.com/thefaylab/hydra_diag

### Data requirements

A common dataset for 10 Georges Bank species has been developed, as well as a simulated dataset for model performance testing. This `mskeyrun` data package holds both datasets. All modeling teams used these datasets. Group decisions on data are also documented below.

**Years:** 1968-2019

**Area:** Georges Bank (see map above)

**Species:** Atlantic cod (*Gadus morhua*), Atlantic herring (*Clupea harengus*), Atlantic mackerel (*Scomber scombrus*), Goosefish (*Lophius americanus*), Haddock (*Melanogrammus aeglefinus*), Silver hake (*Merluccius bilinearis*), Spiny dogfish (*Squalus acanthias*), Winter flounder (*Pseudopleuronectes americanus*), Winter skate (*Leucoraja ocellata*), and Yellowtail flounder (*Limanda ferruginea*)

### Model description

Much of the material below is edited from from Gaichas et al. (2017), supplementary material:

> Hydra is a length-based multispecies, multifleet model designed to provide simulated data for performance testing of simpler (non-size structured) multispecies assessment models and management procedures for the Northeast U.S. Continental Shelf.
>
> We focus on body size rather than age because:
>
> - Predation is a length-based process

- Harvesting is a length-based process
- Routine age determinations are not available for all ecologically and economically important species

**Population dynamics** For an individual species $i$, the number of individuals $N$ in size class $j$ at time $t+1$ represent the survivors from that cohort from the previous time step ($t$) that have not grown into a larger size class, plus the number of individuals growing into size class $j$ from smaller size classes during the time interval. We only consider transitions among immediately adjacent size classes. For size classes larger than the recruit (smallest) size class, the model for each species $i$ can be written:

$$N_{i,j,t+1} = \sum_{j'} \phi_{i,j' \to j,t} S_{i,j',t} N_{i,j',t}$$

The proportion surviving from size class $j-1$ to size class $j$ for a species at time $t$ is denoted $S_{j',t,}$. $\phi_{j' \to j}$ is the probability of moving from size class $j-1$ to size class $j$.

**Growth and Time-in-Stage** The values for $\phi_{j' \to j,t}$, the time required for an individual to grow between length bins are determined from the growth function. Growth can be assumed to follow a von Bertalanffy (see Hall et al. (2006)) or power function, either of which can be applied to any species. As in Hall et al. (2006), the model timestep is set up to equal the amount of time it takes for the fastest growing species/length bin combination to grow into the next length bin.

One of the simplest possible models for indeterminate growth with environmental effects is exponential:

$$l_{i,a,t} = \Psi_i a_i^{K_i}$$

where $l_{i,a,t}$ is the length of species $i$, age $a$ and time $t$; $\Psi_i$ and $K_i$ are growth model parameters.

The time required to grow through a specified length interval $lb$ is:

$$\Delta t_i = \left[ \frac{\Delta lb_i}{\Psi_i} \right]^{1/K_i}$$

and the expression for the probability $\phi$ of moving through the length interval is then:

$$\phi_i = \frac{1}{\Delta t_i}$$

Von Bertalanffy growth is modeled identically as in Hall et al. (2006):

$$l_{i,t} = L_{\infty i}(1 - e^{-vonBK_i,(t - t_{0,i})})$$

with the time required to grow through a specified length interval:

$$\Delta t_i = \frac{1}{vonBK_i} log \left[ \frac{L_{\infty i} - lb_{lower,i}}{L_{\infty i} - lb_{upper,i}} \right]$$

and an identical expression for the probability of moving through the length interval as above.

**Recruitment** Average recruitment with annual deviations

In the most basic formulation, no spawner-recruit model is specified, but annual log deviations $\Lambda_{i,t}$ from mean log recruitment $\bar{R}_i$ determine annual recruitment levels:

$$N_{i,1,t} = e^{\bar{R}_i + \Lambda_{i,t}}$$

**Mortality** The survival rate is:

$$S_{i,j,t} = e^{-(M1_{i,j} + M2_{i,j,t} + s_{i,j,q} F_{i,j,q,t})}$$

The rate of natural mortality is partitioned into two components: sources of natural mortality due to all unmodeled factors ($M1$) and predation mortality from species included in the model ($M2$).

The predation process is decomposed into size selectivity and species vulnerability in each area. The $\rho$, of prey species $m$ size $n$ for a given predator species $i$ size $j$ is:

$$\rho_{i,j,m,n} = \vartheta_{n,j} \lambda_{m,i}$$

where $\vartheta_{n,j}$ is the preference for a prey item of size $n$ by predator size $j$, and $\lambda_{m,i}$ is the vulnerability of prey species $m$ to predator species $i$. The parameter $\lambda$ is set to either 0 or 1 depending on whether predator $i$ is known to prey on species $m$.

The size preference function is:

$$\vartheta_{n,j} = \frac{1}{(w_n/w_j)\sigma_j \sqrt{2\pi}} e^{-\frac{[log_e(w_n/w_j) - \Psi_j]}{2\sigma_j^2}}$$

where $w_n$ is the weight at the midpoint of the length bin for a prey of size $n$, $w_j$ is the weight at the midpoint of the length bin for a predator of size $j$, $\Psi_j$ is the ‘preferred’ predator/prey weight ratio on a logarithmic scale, and $\sigma_j^2$ is the variance in predator size preference.

To estimate the daily food intake $I$ for each predator, we use:

$$I_{i,j,t} = 24[\delta_i e^{\omega_i T}] \bar{C}_{i,j,k,t}$$

where the term in brackets gives the temperature ($T$)-dependent hourly consumption rate with parameters $\delta_i$ and $\omega_i$ and $\bar{C}_{i,j,k,t}$ is the mean stomach content weight (g) over a diel cycle.

The predation mortality rate on a particular prey species depends on the total consumption by predators, the suitability of that prey to each predator, and the total suitable prey biomass available to each of its predators. If the predation functional feeding response is assumed to be a Holling Type II, the $M2$ component on prey species $m$ size $n$ is then (Hall et al., 2006; Magnússon, 1995):

$$M2_{m,n,t} = \sum_i \sum_j I_{i,j,t} N_{i,j,t} \frac{\rho_{i,j,m,n}}{\sum_a \sum_b \rho_{i,j,a,b} W_{a,b} N_{a,b} + \Omega}$$

where $a$ and $b$ represent all prey species and sizes for predator $i$, $W$ is the mean weight of prey $a$ in size class $b$, $\Omega$ is ‘other’ food not explicitly included in the model, and all other terms are as above.

Fishing in the hydra estimation model has been re-implemented differently from the original simulation model. Similar to the original simuilation model, multispecies fleets are implemented. However, rather than driving a simulation with fleet effort, we estimate fishing mortality rates ($F$) by fleet where there are one or more target species/stocks per fleet, with one “primary” stock identified for each fleet.

Annual fishing mortality rates ($F$) over time for the primary stock (modeled via average $Fbar$ and annual deviations $Fdev$) are specified as:

$$F_{f,i,y} = Fbar_{f,i}e^{(Fdev_{f,y})}$$

fleet-specific $F$ for target stocks other than the primary stock modeled as offset from the primary stock via catchability $q$:

$$F_{f,j,y,} = q_{f,j}.F_{f,i,y}$$

We note this implies same fishing pattern across stocks within a fleet, which is a simplification that may need later adjustment. This could be modified with a multivariate random walk.

Logistic selectivity at length for each fleet is applied to all species caught by the fleet:

$$F_{f,i,t,l} = sel_{f,i,L(l)}$$

The selectivity for a specified gear type $q$ is given by:

$$s_{i,j,q} = [1 + e^{-(c_{i,q}+d_{i,q}l)}]^{-1}$$

where $c_{i,j,q}$ and $d_{i,j,q}$ are model parameters and $l$ is the midpoint of the $j$th length interval.

The yield from species $i$ at time $t$ is then:

$$Y_{i,t} = \sum_j \sum_q \frac{s_{i,j,q}F_{i,j,q,t}}{M1_{i,j} + M2_{i,j,t} + s_{i,j,q}F_{i,j,q,t}}[1 - e^{-(M1_{i,j}+M2_{i,j,t}+s_{i,j,q}F_{i,j,q,t})}]N_{i,j,t}W_{i,j}$$

where $W_{i,j,k}$ is the mean weight of an individual of species $i$ and length $j$.

### Objective Function

The model can be fitted to multiple data streams. Currently the objective function comprises the following components, catch (by species for each fleet), size composition of the catch (by species for each fleet), survey abundance index (by species for each survey), survery size composition (by species for each survey), survey diet composition (stomach weight by predator size bin for each predator species and survey), and recruitment (annual deviations from mean recruitment by species).

**Catch** Expected catch for species $i$ in fleet $q$ is summed over time steps to the fishing/calendar year $y$.

$$\hat{C}_{y,i,q} = \sum_{t \in y} \sum_j \frac{s_{i,j,q}F_{i,j,q,t}}{M1_{i,j,t} + M2_{i,j,t} + s_{i,j,q}F_{i,j,q,t}} \left[1 - e^{-(M1_{i,j,t}+M2_{i,j,t}+s_{i,j,q}F_{i,j,q,t})}\right] N_{i,j,t}W_{i,j,t}$$

$$ln(C_{y,i,q}) \sim \mathcal{N}(ln(\hat{C}_{y,i,q}), \sigma^2_{q,i,y})$$

### Catch length composition

$$\hat{p}_{q,i,y,j} = p^*_{q,i,y,j} / \sum_{j'} p^*_{q,i,y,j'}$$

$$p^*_{q,i,y,j} = \sum_{t \in y} \frac{s_{i,j,q}F_{i,j,q,t}}{M1_{i,j,t} + M2_{i,j,t} + s_{i,j,q}F_{i,j,q,t}} \left[1 - e^{-(M1_{i,j,t}+M2_{i,j,t}+s_{i,j,q}F_{i,j,q,t})}\right] N_{i,j,t}$$

Proportions at length (in numbers) in catch for each species in each fleet are modeled as multinomial, given an input annual sample size for each composition.

**Survey abundance indices** The predicted values for survey abundance rely on the timing of the survey within a year, catchability of the species in the survey, and the (length-based) survey selectivity. Survey $g$ predicted values (in weight) are:

$$\hat{I}_{g,i,t} = q_{i,g,t} \sum j s_{i,j,g} N_{i,j,t} l_j W_{i,j,t}$$

The observed annual index values for each species are assumed to be lognormally distributed around the predicted values:

$$ln(I_{g,i,t}) \sim \mathcal{N}(ln(\hat{I}_{g,i,t}), \sigma^2_{g,i,t})$$

**Survey length composition**

$$\hat{p}_{g,i,t,j} = p^*_{g,i,t,j} / \sum_{j'} p^*_{g,i,t,j'}$$

$$p^*_{g,i,t,j} = s_{i,j,g,t} N_{i,j,t}$$

Proportions at length (in numbers) in the survey for each species are modeled as multinomial, with an additional input sample size for each composition.

**Prey proportions in survey diet** Proportions by weight in stomach of predator $i$ of size $j$ of prey $m$:

$$\hat{\phi}_{g,t,i,j,m} = \phi^*_{g,t,i,j,m} / \left( \phi^*_{g,t,i,j,other} + \sum_m \phi^*_{g,t,i,j,m} \right)$$

$$\phi^*_{g,t,i,j,m} = \sum_n \rho_{i,j,m,n} N_{m,n,t} W_{m,n,t}$$

Proportions by prey species (in weight) in the survey are modeled as multinomial with an input sample size for each composition (potentially these data could be modeled as Dirichlet or delta-dirichlet, or....)

### Penalties

**Recruitment deviations** A penalty to the objective function is added for the annual deviations from the stock-recruitment relationship for each species, which are assumed to be log-normal around the expected recruitment (equation 8):

$$\Lambda_{i,t} \sim \mathcal{N}\left(0, \tau^2_{\Lambda_i}\right)$$

where $\tau^2_{\Lambda_i}$ is the variance of the recruitment deviations for species $i$.

### Estimated (or estimable) parameters

Parameter types and numbers of parameters by type (Nsp = number of species, Nlen = number of length bins, Nyr = number of years, Nfleet = number of fleets, Nsurvey = number of surveys) include:

1. Initial year numbers at size for each stock (Nsp * Nlen)

2. Average annual recruits for each stock (Nsp)

3. Recruitment deviations for each stock (Nsp * (Nyr-1) )

4. Annual F for each fleet (Nfleet * Nyr)

5. Fishery catchability for non-primary stocks (Nsp - Nfleet)

6. Fishery logistic selectivity at length for each fleet (2 * Nfleet)

7. Survey catchability for each stock for each survey (Nsurvey * Nsp)

8. Survey logistic selectivity at length for each survey (2 * Nsurvey)

Not currently estimating (but could add):

1. M1 by stock (constant over age) (Nstock)

2. M2 related parameters (e.g. vulnerability)

### Diagnostics

Functions to visualize both model inputs and outputs are in development in the hydra_diag github repository. At present, workflows for visualizing an individual model output and for comparing across model outputs are available, along with initial options to add model skill assessment (comparison with known simulated data).

Individual model diagnostics include plots of input survey biomass, catch, and length data (not aggregated by length bin). Model fits to aggregate survey indices and catch by fleet, fits to annual and aggregate (all years) survey and fishery length compositions (length bins), and diet composition for predator and length bin are visualized. In addition, effective sample size is compared with input sample size for length and diet compositions, and both Pearson and One Step Ahead (OSA, Trijoulet et al. (2023)) residuals are calculated for fits to all length compositions. Model derived quantities (estimated recruitment, biomass, fishing mortality, predation mortality, and survey and fishery selectivity) are also plotted.

### Modeling approach

Given the current development level of Hydra as an estimation model, current objectives are to demonstrate that the model runs and generates the desired output. Priorities for further development should ideally be identified during the ECES WGSAM review. Hydra is being fit to both simulated and real Georges Bank datasets to develop and test estimation functions.

For simulated data our initial species selection includes 11 single species groups from the Norwegian Barents Sea (NOBA) Atlantis model (Hansen et al. 2016, 2019). These groups are fully age structured. All but two of them are fished. The full process for generating the simulated dataset is described on the `mskeyrun` R package documentation at this link. Hydra was fit to an 80 years of simulated survey and catch index, size, and diet composition data for these 11 simulated stocks using 2 fleets (cod and other), with both 5 and 10 length bins for each stock. The simulated datasets were used to test incrementally added estimation functionality, and will be used for model skill assessment.

For Georges Bank, the model was fit to 10 stocks using 2 fleets (groundfish and pelagic), with 5 length bins for each stock. Time series for fitting included 1978 - 2019. The model was fit to total catch (summed retained + discard), catch size compositions, annual Fall and Spring NEFSC bottom trawl survey abundance indices (pre-processed for vessel differences via calibration), survey size compositions, survey and stomach content diet composition (proportion by weight). M1 and M2-related mortality parameters are fixed inputs (for now).

Initial fits to real Georges Bank data have been for testing only, but several examples were presented during the review. Please see the “Hydra Results” document for an overview.

# References

Fournier, David A., Hans J. Skaug, Johnoel Ancheta, James Ianelli, Arni Magnusson, Mark N. Maunder, Anders Nielsen, and John Sibert. 2012. "AD Model Builder: Using Automatic Differentiation for Statistical Inference of Highly Parameterized Complex Nonlinear Models." *Optimization Methods and Software* 27 (2): 233–49. https://doi.org/10.1080/10556788.2011.597854.

Gaichas, Sarah K., Michael Fogarty, Gavin Fay, Robert Gamble, Sean Lucey, and Laurel Smith. 2017. "Combining Stock, Multispecies, and Ecosystem Level Fishery Objectives Within an Operational Management Procedure: Simulations to Start the Conversation." *ICES Journal of Marine Science* 74 (2): 552–65. https://doi.org/10.1093/icesjms/fsw119.

Hall, Stephen J, Jeremy S Collie, Daniel E Duplisea, Simon Jennings, Mark Bravington, and Jason Link. 2006. "A Length-Based Multispecies Model for Evaluating Community Responses to Fishing." *Canadian Journal of Fisheries and Aquatic Sciences* 63 (6): 1344–59. https://doi.org/10.1139/f06-039.

Hansen, Cecilie, Kenneth F. Drinkwater, Anne Jähkel, Elizabeth A. Fulton, Rebecca Gorton, and Mette Skern-Mauritzen. 2019. "Sensitivity of the Norwegian and Barents Sea Atlantis End-to-End Ecosystem Model to Parameter Perturbations of Key Species." Edited by Athanassios C. Tsikliras. *PLOS ONE* 14 (2): e0210419. https://doi.org/10.1371/journal.pone.0210419.

Hansen, Cecilie, Mette Skern-Mauritzen, Gro van der Meeren, Anne Jähkel, and Kenneth F. Drinkwater. 2016. *Set-up of the Nordic and Barents Seas (NoBa) Atlantis Model.* Fisken Og Havet;2-2016. Havforskningsinstituttet. https://imr.brage.unit.no/imr-xmlui/bitstream/handle/11250/2408609/FoH_2-2016.pdf?sequence=1&isAllowed=y.

Rochet, Marie-Joëlle, Jeremy S. Collie, Simon Jennings, and Stephen J. Hall. 2011. "Does Selective Fishing Conserve Community Biodiversity? Predictions from a Length-Based Multispecies Model." *Canadian Journal of Fisheries and Aquatic Sciences* 68 (3): 469–86. https://doi.org/10.1139/F10-159.

Trijoulet, Vanessa, Christoffer Moesgaard Albertsen, Kasper Kristensen, Christopher M. Legault, Timothy J. Miller, and Anders Nielsen. 2023. "Model Validation for Compositional Data in Stock Assessment Models: Calculating Residuals with Correct Properties." *Fisheries Research* 257 (January): 106487. https://doi.org/10.1016/j.fishres.2022.106487.

# Hydra Results, WGSAM Review

Gavin Fay, Sarah Gaichas, Cristina Perez

October 2022

## Introduction

Hydra model equations, new objective function equations, and descriptions of initial modeling approaches and diagnostics are found in the Hydra Methodology document. An overview of results presented during the ICES WGSAM 2022 meeting under ToR b, keyrun reviews, is included in this document.

Hydra model diagnostics and progress fitting to three datasets were presented:

- Simulated data with 5 length bins
- Simulated data with 10 length bins
- Georges Bank data with 5 length bins

We use 5 length bins as a starting structure for fitting the Georges Bank data because the simulation model Hydra used 5 length bins. We do not consider any of these models fully converged as keyruns, rather the diagnostics with each dataset are intended to give insight into model behavior and performance.

Specifically, the simulated datasets with different numbers of length bins are intended to evaluate model sensitivity to this choice of input data structure and tradeoffs with overall performance. These datasets also allow preliminary comparisons with true Atlantis quantities (biomass, recruitment, diet) as a step towards skill assessment. Sensitivity of the model to the Other food parameter was identified as a high priority in the 2018 simulation model review, and is being evaluated with the simulated dataset. Initial fits to different levels of other food with simulated data were posted, but not presented at the meeting due to time constraints.

Fits to the real Georges Bank data have just begun, and have already highlighted several implicit assumptions in the Hydra model code that have been addressed. The time period for modeling has been adjusted to 1978-2019 to avoid starting the model with extremely high catches observed 1968-1977.

Three sets of Georges Bank fits were presented, including:

- Estimate only average F, no initial N or recruitment deviations
- Estimate average F and initial N, no recruitment deviations
- Estimate all

## Results

### These fits are shown to illustrate initial model performance. Neither the model code nor the data inputs are considered final.

Overall, the results of initial fitting are encouraging, considering that Hydra has only recently been developed as an estimation model, and the datasets continue to be refined as well. We note that interannual variability has been picked up by the model for lots of stocks in both simulated and real datasets, suggesting that the model is able to respond to input data. The fits and other outputs shown here are from the model "estimate-all" presented 11 October 2022.

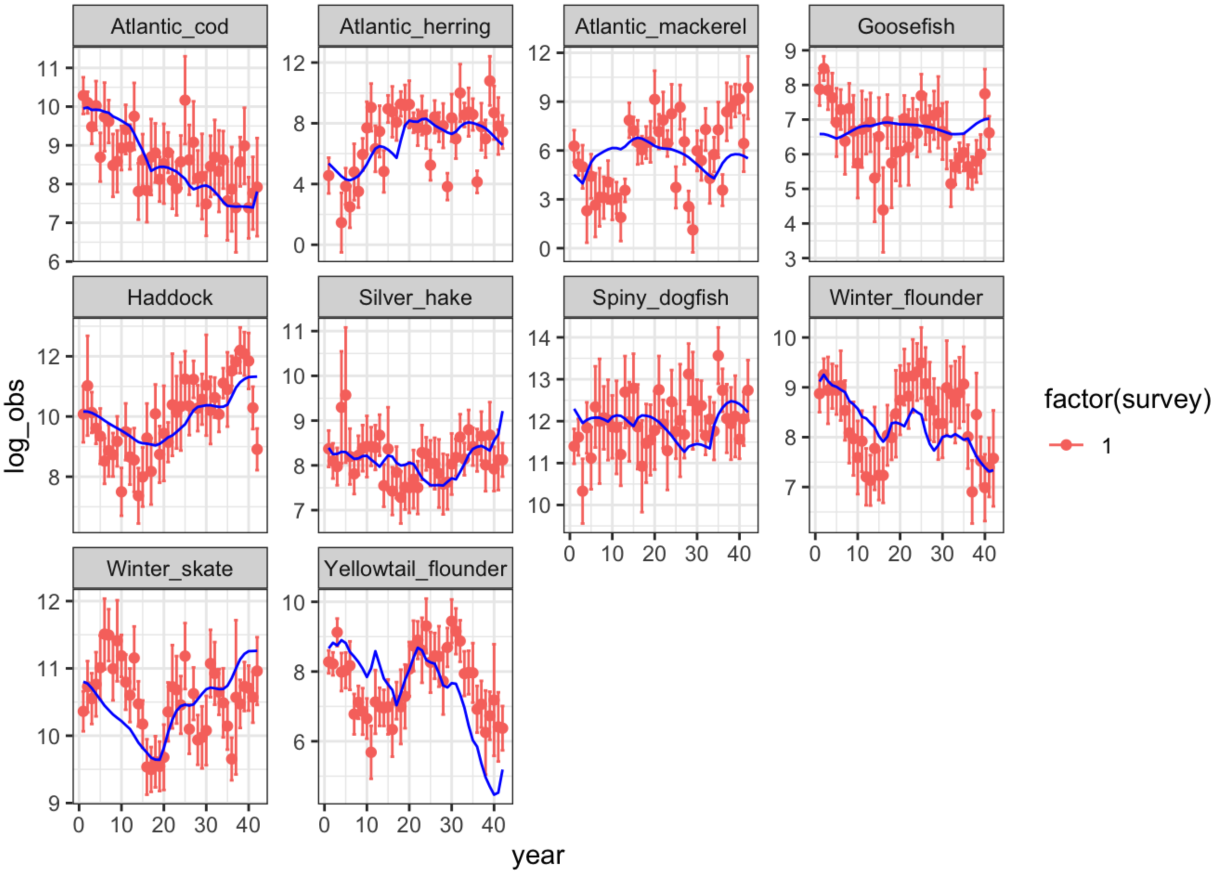


Figure 1: Survey 1index fit, Georges Bank estimate-everything Oct 11 2022

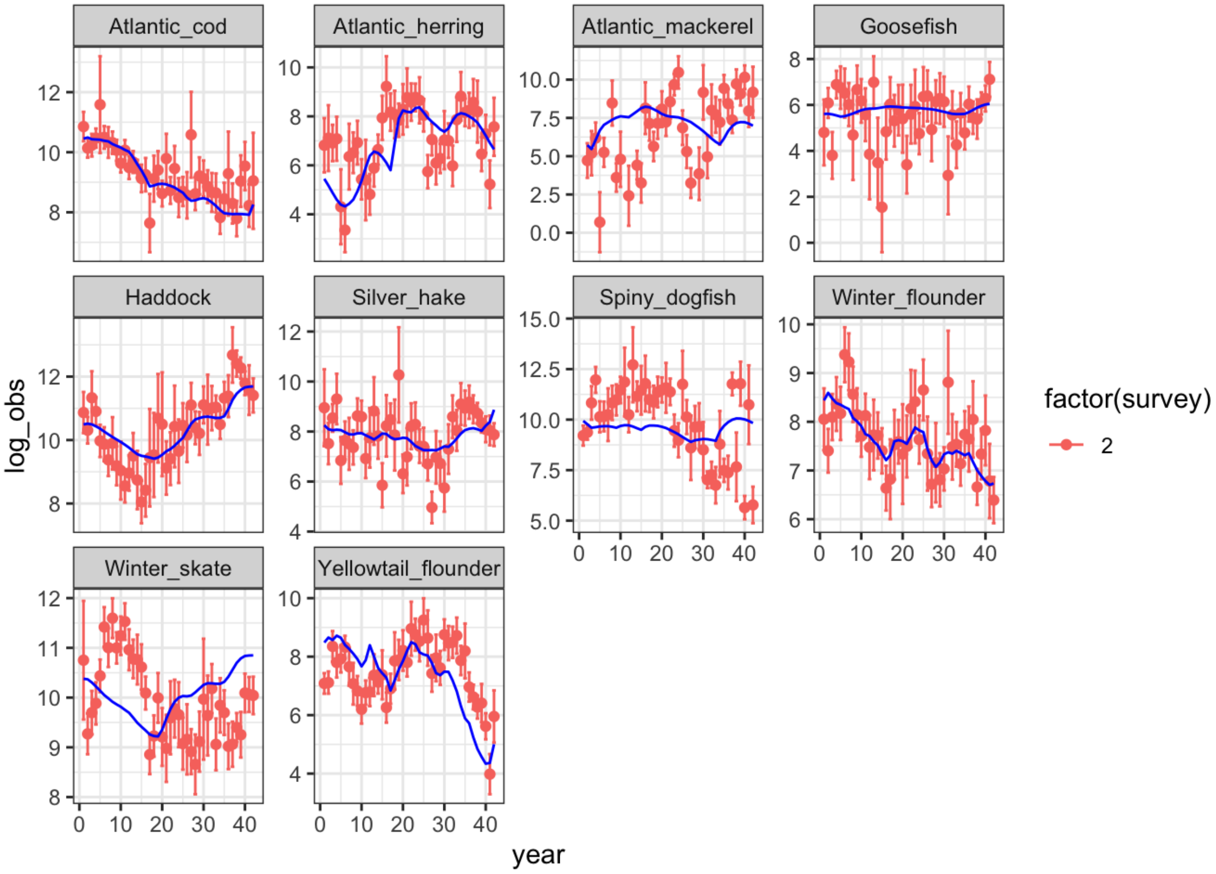


Figure 2: Survey 2 index fit, Georges Bank estimate-everything Oct 11 2022

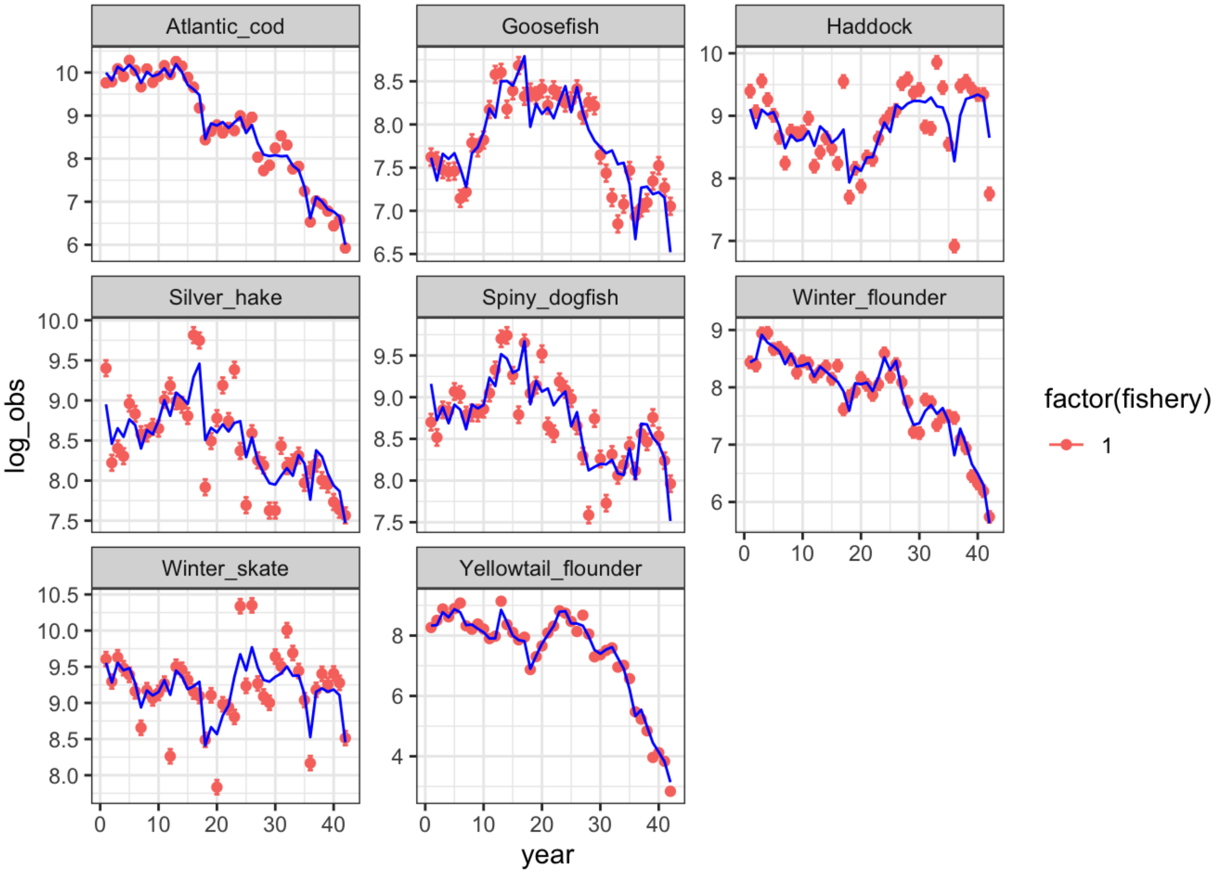


Figure 3: Fleet 1 catch fit, Georges Bank estimate-everything Oct 11 2022

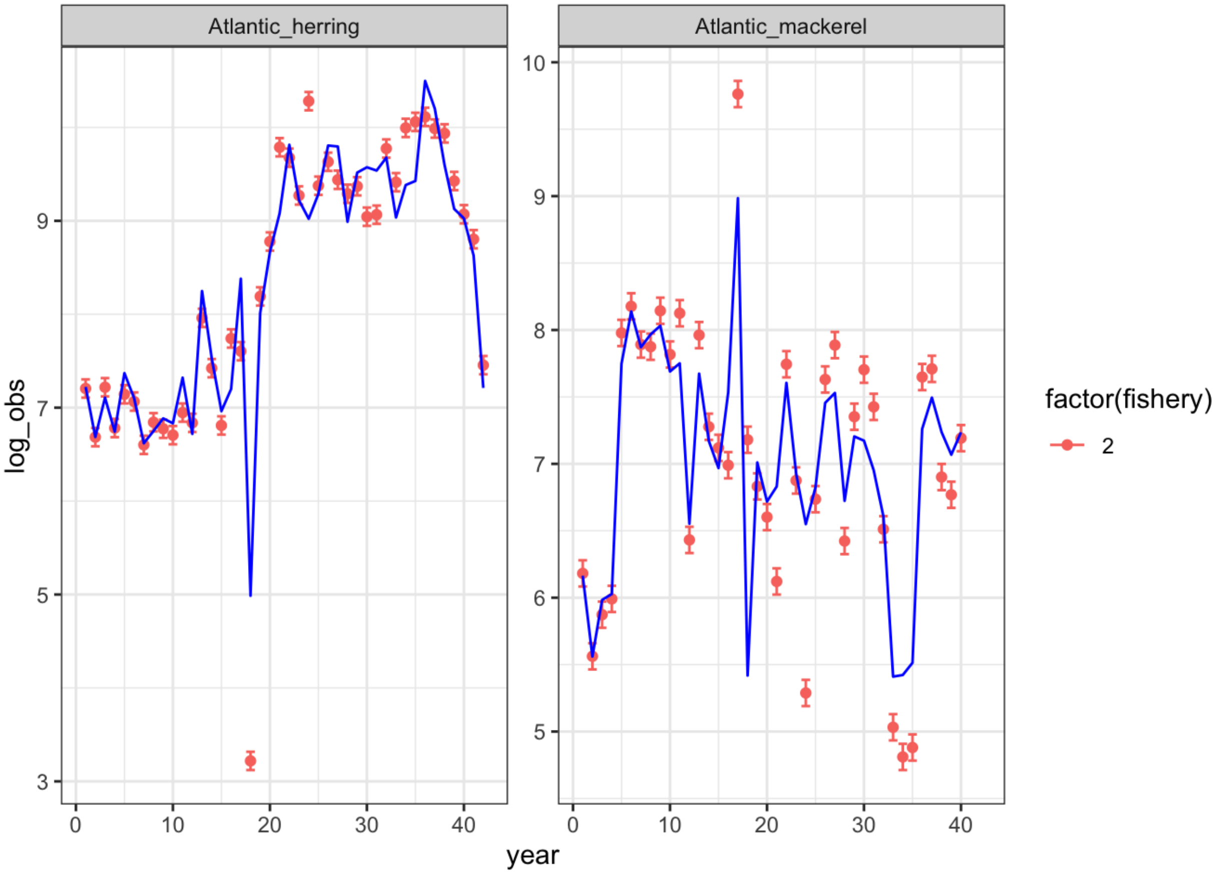


Figure 4: Fleet 2 catch fit, Georges Bank estimate-everything Oct 11 2022

However, it is clear that there is a challenge estimating scale for some stocks. In particular, Georges Bank pelagics (Atlantic herring and Atlantic mackerel), scale is problematic, and estimated F is very low. This is also true for yellowtail flounder (notably, single stock assessment for this stock on Georges Bank is difficult as well).

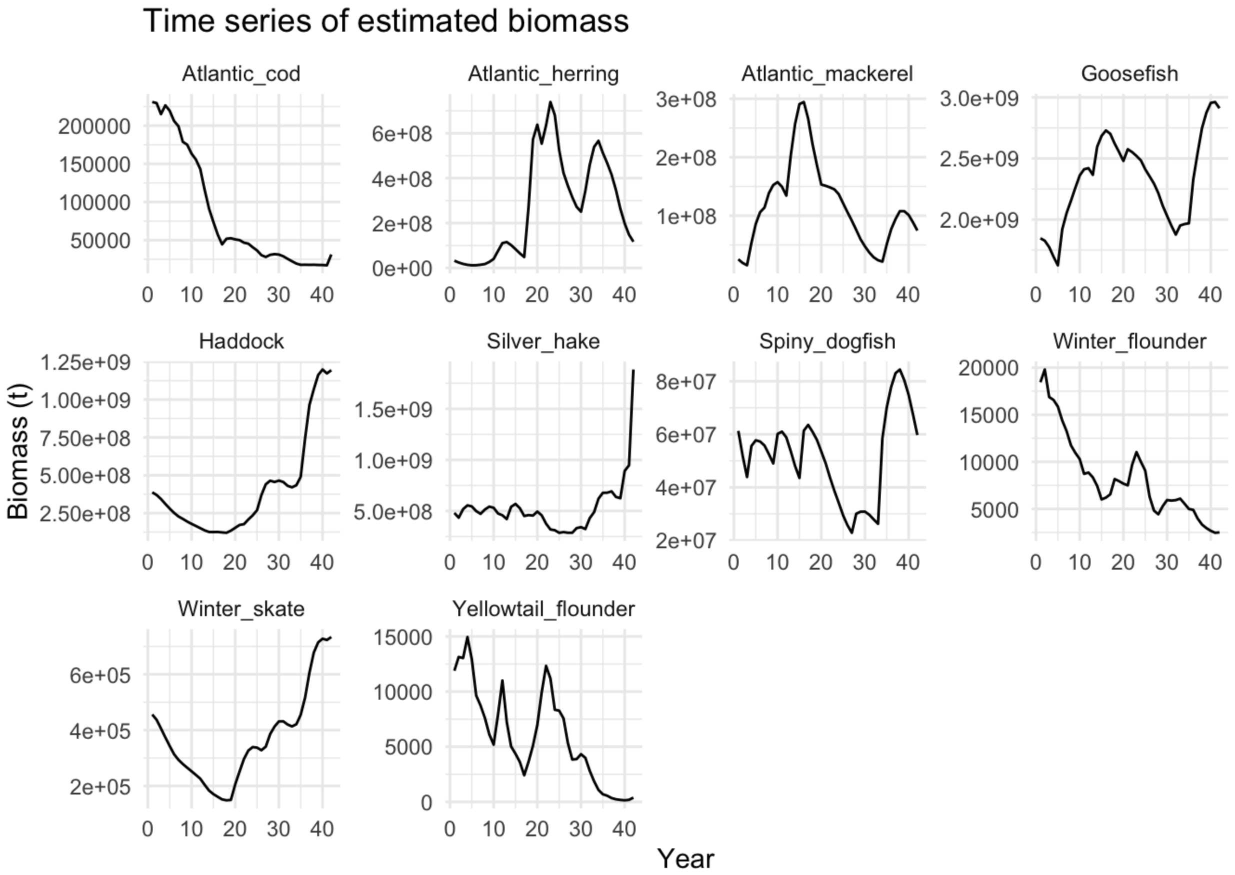


Figure 5: Hydra estimated biomass, Georges Bank estimate-everything Oct 11 2022

Fits to annual size composition in these initial attempts are poor, and turning off diet composition during fitting doesn't affect general model behavior too much.

Model estimated recruitment deviations show patterns that are unlikely given species life history (e.g. spiny dogfish).

To begin addressing where fitting issues are happening, a series of runs are in progress with Georges Bank data sequentially turning on different portions of the estimation. In addition, fits to simulated data are in progress to determine the impacts of different length bin specification on model estimation performance. Finally, sensitivity analysis has started looking at different levels of OtherFood using simulated datasets. Results from all of these analyses in progress were available during the review.

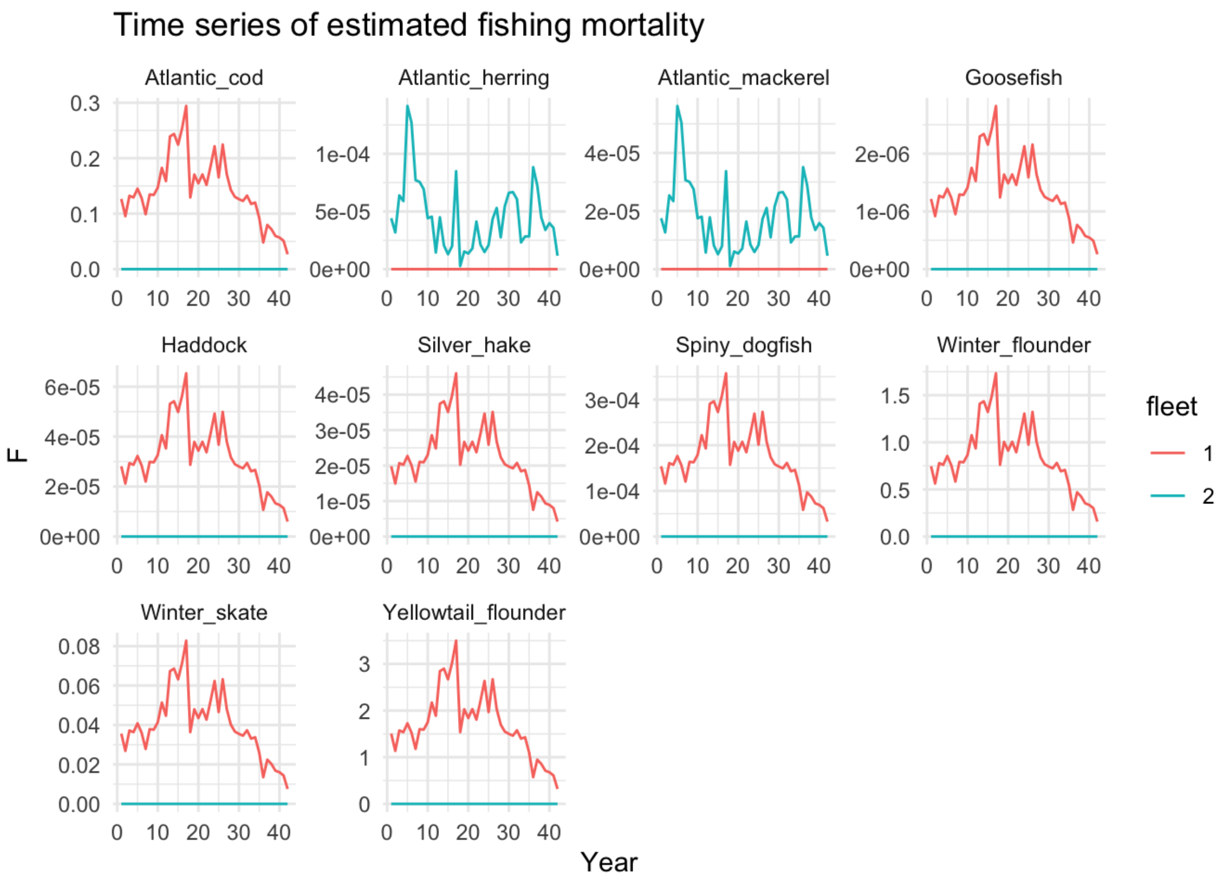


Figure 6: Hydra estimated F, Georges Bank estimate-everything Oct 11 2022

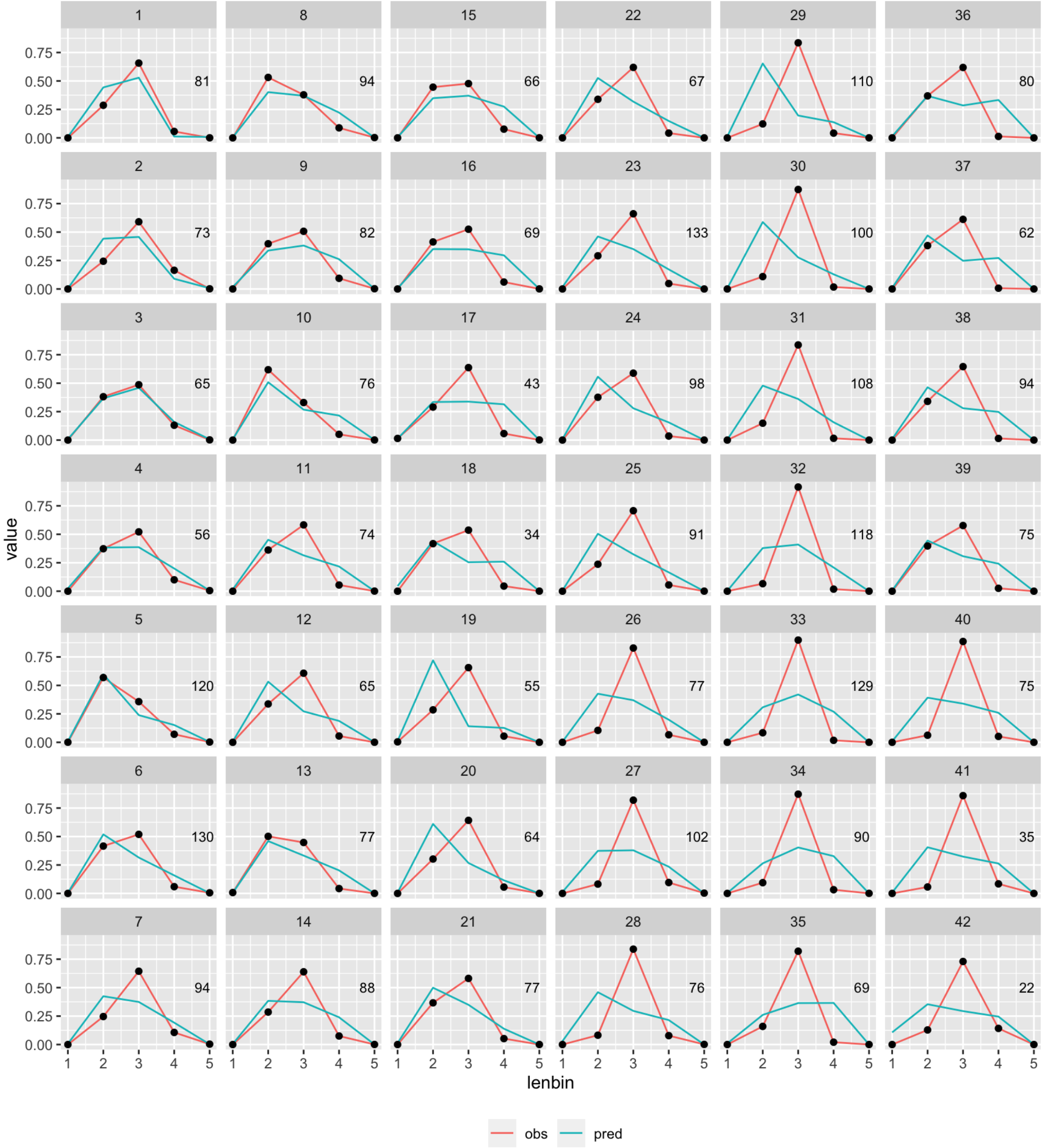


Figure 7: Atlantic cod fishery length comp fit, Georges Bank estimate-everything Oct 11 2022. This example shows fits to annual Atlantic cod fishery length compositions (in the 5 model length bins, each panel represents a year).

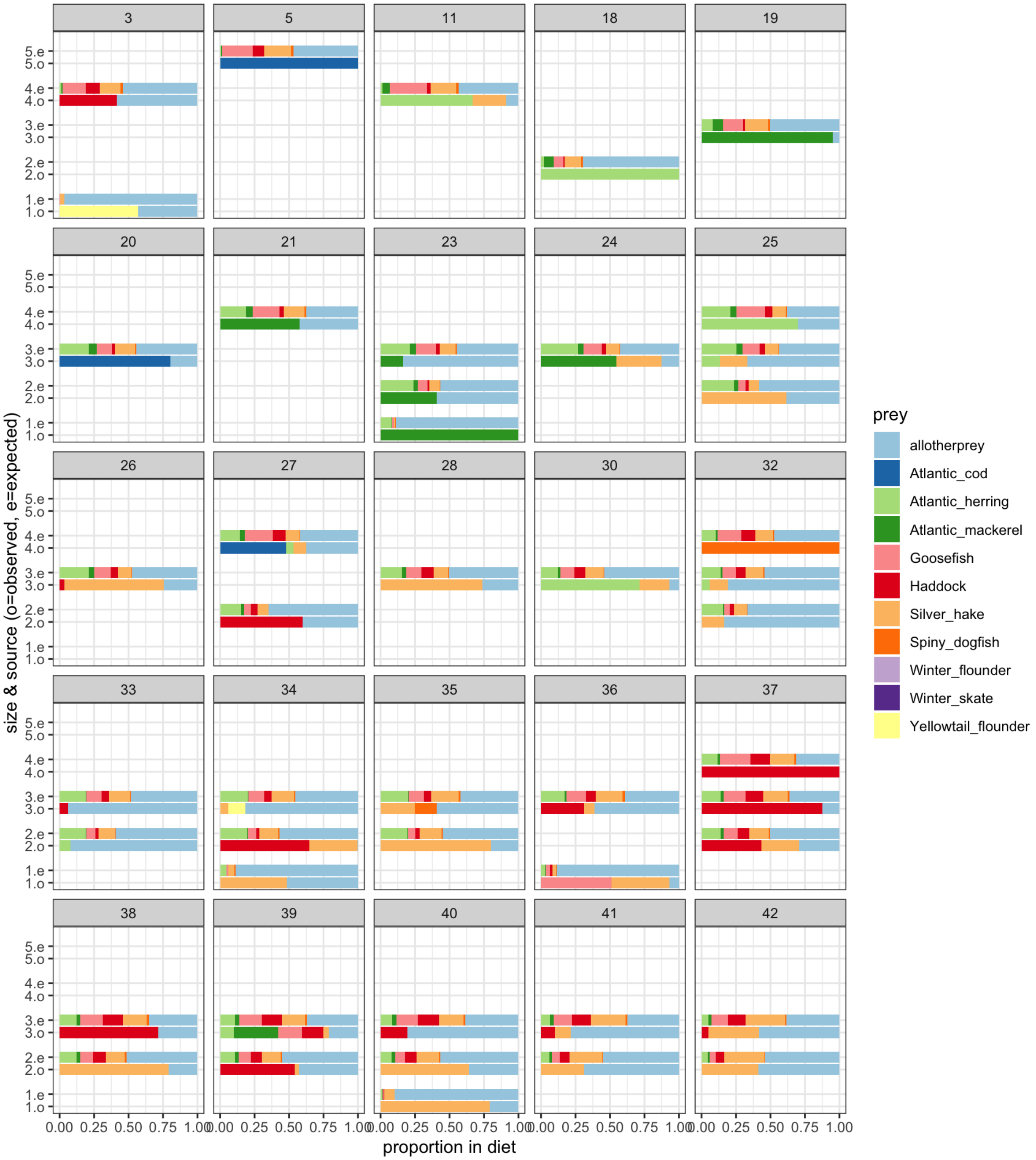


Figure 8: Monkfish diet comp fit, Georges Bank estimate-everything Oct 11 2022. This example compares observed and estimated diet composition at length for Monkfish (Fall survey, each panel represents a year).

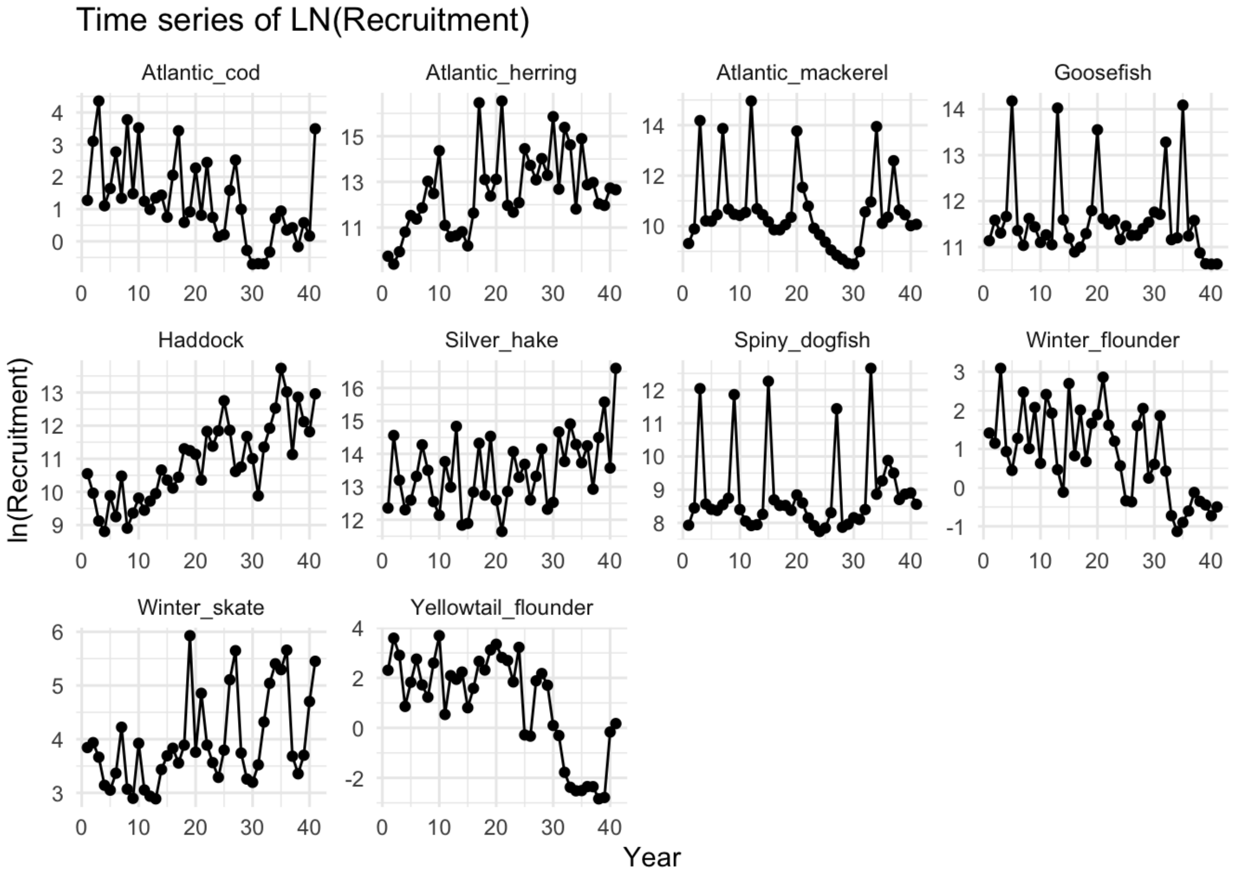


Figure 9: Hydra estimated recruitment, Georges Bank estimate-everything Oct 11 2022

## Next steps

Initial skill assessment and initial fits to GB mskeyrun data suggest Hydra is responsive to data sets. However, ongoing modeling work is designed to get more of a feel for how this model is behaving when confronted with data.

Potential topics for discussion with reviewers include:

- Mismatch between drivers of trends and spatial resolution (subsets of stock being included for pelagics)
- Generic structure of length bin parameterization
- Fishing fleet specification
- Fits to all species together rather than starting from results of single-species
- Predation mortality component estimation
- Data weighting
- Simulation testing (self-tests, fits to Atlantis)
- Move to TMB (state space) framework (Rceattle?)
- Application as a multispecies assessment model
- How precise do the growth functions need to be? (i.e. reliance on age/growth studies)
- How do growth model choices (structure) affect performance

# Annex 5: Rpath Model

## A mass balance representation of Georges Bank, USA

Sean M. Lucey[1]

[1]NOAA, National Marine Fisheries Service, Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543

**ARTICLE HISTORY**
Compiled January 9, 2023

**ABSTRACT**
Georges Bank is a highly productive region of the Northeast US Continental Shelf Large Marine Ecosystem. Its unique physical properties make it ideal for ecosystem-based fisheries management which is inherently place-based. In order to evaluate ecosystem-based strategies there will need to be a management strategy evaluation procedure undertaken. As part of this procedure various operating models will need to be developed to account for the various uncertainty in the system. Mass balance models offer many of the characteristics that would benefit an ecosystem-level management strategy evaluation. Existing mass balance models in the region are either too aggregated at both the species and fleet level or not geographically specified to Georges Bank. Here we describe a new mass balance model for Georges Bank built using Rpath, the R implementation of the mass balance algorithms. Our model has 71 groups including 69 living groups, 2 detrital groups, and 10 fleets. It describes a system that is highly productive and inter-connected. The package in which it was built, Rpath, is flexible and should allow for the model to be tailored to address specific management questions, an important feature as management timeline tend to be more accelerated than model development horizons.

**KEYWORDS**
Rpath; mass balance; Georges Bank; EBFM

## 1. Introduction

Included in the suite of MSKeyrun models is a mass balance representation of Georges Bank. Mass balance models are a good tool for identifying and quantifying major energy flows in a system (Plaganyi 2007). Mass balance models describe the ecosystem resources and their interactions which makes them good at evaluating the ecosystem effects of fishing or environmental changes as well as exploring management policy options (Plaganyi 2007). The use of mass balance models was popularized through the use of Ecopath with Ecosim (EwE: Christensen and Pauly 1992; Walters, Christensen, and Pauly 1997; Christensen and Walters 2004). For MSKeyrun, I used the R version, Rpath (Lucey, Gaichas, and Aydin 2020).

Mass balance models can be used in a variety of ways. For the ICES WGSAM review, only the initial mass balance was complete. Mass balance is a static snapshot of the ecosystem and is a trophic model at its base (Polovina 1984; Christensen and Pauly 1992, Figure 1). Having a balanced model does not necessarily imply steady-

---

CONTACT Sean M. Lucey. Email: `sean.lucey@noaa.gov`

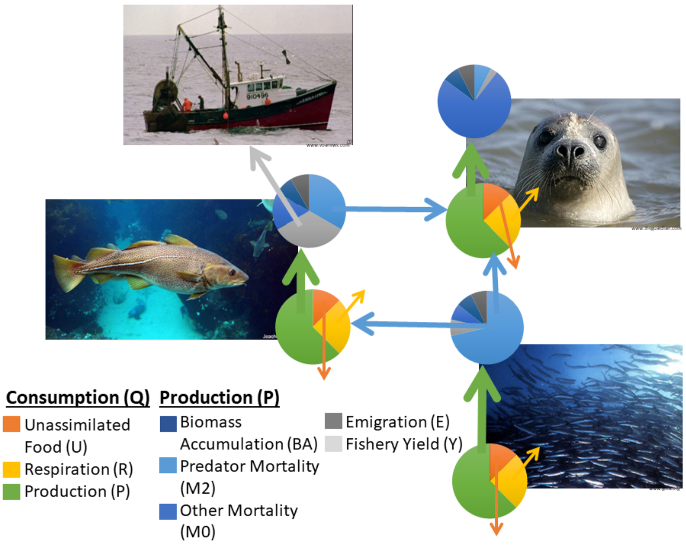


**Figure 1.** Conceptual model of mass balance where consumption and production are balanced. Production as described in Equation 1 as a sum of fishery yield (Y), predator mortality (M2), emigration (E), biomass accumulation (BA), and other mortality (M0 or 1 − EE). Consumption as described in Equation 2 is a sum of production (P), unassimilated food (U) and respiration (R).

state, although many models are parametrized as such. Mass balance models can also be extended into dynamic simulations referred to as Ecosim in the EwE framework or Rsim in the Rpath framework (Walters, Christensen, and Pauly 1997; Lucey, Gaichas, and Aydin 2020). Spatial processes can also be explored using the Ecospace module of EwE (Walters, Pauly, and Christensen 2000). No such equivalent exists in Rpath at this time. Future work within the MSKeyrun project will include dynamic simulations as well as tuning to time series data.

This new mass balance representation of Georges Bank build off existing mass balance work that was developed as part of the Northeast Fisheries Science Center's (NEFSC) Energy Modeling and Analysis eXercise (EMAX: Link et al. 2006). EMAX models were developed for both the Ecopath (Christensen and Pauly 1992; Walters, Christensen, and Pauly 1997; Christensen and Walters 2004) and EcoNetwrk (Ulanowicz and Kay 1991) software packages. The Georges Bank model from EMAX is highly aggregated with 29 species groups, two detrital groups, and one fishery. This trophic structure is too coarse to be useful for management. Many of the commercially important species are aggregated together and there is not enough resolution in the fisheries to address any meaningful trade-offs.

A companion model from EMAX describing the Gulf of Maine was incorporated into the early stages of a management strategy evaluation (MSE) on the role of Atlantic herring, *Clupea harengus* as a forage item within the system particularly for marine mammals. Despite EMAX being designed to describe forage fish in the region (Link et al. 2009), the results were highly uncertain with no discernible benefits due to the highly aggregated nature of the model. Therefore the model was not pursued further within MSE (S. Gaichas personal communication).

The purpose of the MSKeyrun project is to make models that will be useful for managers. With that goal in mind, I present a new mass balance representation of Georges Bank. Here all commercial species have been disaggregated and the fisheries have been divided into major gear categories. Developing the model in Rpath will also let us explore various fisheries management strategies using built in functionality (Lucey et al. 2021).

## 2. Methods

### 2.1. *Mass Balance Equations*

Mass balance is governed by two master equations; one describing production (Eq. 1) and the other consumption (Eq. 2: Christensen and Pauly 1992). The interaction between the two gives rise to the term mass balance. Production accounts for natural and fishery-induced mortality as well as emigration and biomass accumulation as:

$$P_i = Y_i + B_i M2_i + E_i + BA_i + P_i(1 - EE_i) \tag{1}$$

where production, $P_i$, is the sum of a species' fishery yield, $Y_i$; predation mortality, $M2_i$; emigration, $E_i$; biomass accumulation, $BA_i$; and other mortality, expressed as $P_i(1 - EE_i)$. Where $EE_i$ is the ecotrophic efficiency or percentage of mortality explained within the model. In this equation, $M2_i$ is expressed as a rate and therefore multiplied by the species biomass, $B_i$.

Consumption accounts for the production expressed above as well as unassimilated food and respiration as:

$$Q_i = P_i + R_i + U_i \tag{2}$$

where $Q_i$ is the total consumption, $R_i$ is respiration, and $U_i$ is unassimilated food. Energy used for respiration is lost from the system. Unassimilated food is the portion of consumption that is excreted and remains in the system through a detrital group.

### 2.2. *Parameter input*

Mass balance models require several parameters that can be obtained from commonly collected data. These parameters include biomass, catch, and food habits. In addition, several biological parameters that relate biomass to production and consumption are required. The following sections will outline how inputs were derived.

#### *2.2.1. Fisheries independent data*

Many fish species as well as large benthic animals are routinely sampled by the NEFSC (Table 1). For the purpose of this model the Autumn Bottom Trawl Survey (Politis et al. 2014), the Sea Scallop and Integrated Benthic Survey (NEFSC 2018), and the Atlantic Surf Clam and Ocean Quahog Survey (Jacobson and Hennen 2019) provided biomass estimates. Data for years 1981 to 1985 were queried from the NEFSC survey databases. This time period was chosen to mimic models for the Gulf of Maine and Mid-Atlantic Bight that are also under development.

Table 1.: Living groups for the Georges Bank Rpath model whose data were pulled directly from Northeast Fishery Science Center (NEFSC) databases. Surveys used are the Autumn Bottom Trawl Survey (ABTS), Sea Scallop and Integrated Benthic Survey (Scallop), and the Atlantic Surf Clam and Ocean Quahog Survey (Clam). Classifications and diet are those used for PREBAL calculations.

| Group | Data Source | Classification | Diet |
|---|---|---|---|
| AtlHerring | ABTS | Pelagic (Small; Round) | Planktivore |
| AtlMackerel | ABTS | Pelagic (Small; Round) | Planktivore |
| Butterfish | ABTS | Pelagic (Small; Round) | Benthivore |
| SmPelagics | ABTS | Pelagic (Small; Round) | Planktivore |
| Mesopelagics | ABTS | Pelagic (Small; Round) | Benthivore |
| OtherPelagics | ABTS | Pelagic (Medium; Round) | Piscivore |
| Cod | ABTS | Demersal (Round) | Piscivore |
| Haddock | ABTS | Demersal (Round) | Benthivore |
| Goosefish | ABTS | Demersal | Piscivore |
| OffHake | ABTS | Demersal (Round) | Piscivore |
| SilverHake | ABTS | Demersal (Round) | Piscivore |
| RedHake | ABTS | Demersal (Round) | Benthivore |
| WhiteHake | ABTS | Demersal (Round) | Benthivore |
| Redfish | ABTS | Demersal (Round) | Benthivore |
| Pollock | ABTS | Demersal (Round) | Benthivore |
| OceanPout | ABTS | Demersal | Benthivore |
| BlackSeaBass | ABTS | Demersal (round) | Benthivore |
| Bluefish | ABTS | Pelagic (Medium; Round) | Piscivore |
| Scup | ABTS | Pelagic (Small; Round) | Benthivore |
| OtherDemersals | ABTS | Demersal (Round) | Benthivore |
| SouthernDemersals | ABTS | Demersal (round) | Benthivore |
| Fourspot | ABTS | Demersal (Flat) | Piscivore |
| SummerFlounder | ABTS | Demersal (Flat) | Piscivore |
| AmPlaice | ABTS | Demersal (Flat) | Benthivore |
| Windowpane | ABTS | Demersal (Flat) | Benthivore |
| WinterFlounder | ABTS | Demersal (Flat) | Benthivore |
| WitchFlounder | ABTS | Demersal (Flat) | Benthivore |
| YTFlounder | ABTS | Demersal (Flat) | Benthivore |
| OtherFlatfish | ABTS | Demersal (Flat) | Benthivore |
| SmFlatfishes | ABTS | Demersal (Flat) | Benthivore |
| SpinyDogfish | ABTS | Demersal (Round) | Piscivore |
| SmoothDogfish | ABTS | Demersal (Round) | Benthivore |
| Barndoor | ABTS | Demersal (Flat) | Piscivore |
| WinterSkate | ABTS | Demersal (Flat) | Benthivore |
| LittleSkate | ABTS | Demersal (Flat) | Benthivore |
| OtherSkates | ABTS | Demersal (Flat) | Benthivore |
| Illex | ABTS | Invertebrate (Pelagic) | |
| Loligo | ABTS | Invertebrate (Pelagic) | |
| OtherCephalopods | ABTS | Invertebrate (Pelagic) | |
| AmLobster | ABTS | Invertebrate (Benthic) | |
| AtlScallop | Scallop | Invertebrate (Benthic) | |
| Clams | Clam | Invertebrate (Benthic) | |

Samples pertaining to Georges Bank were identified using the same stratification as used for the NEFSC State of the Ecosystem reports (Figure 2) Dominate or commercially important species were kept as separate species groups while less dominate/important species were grouped into aggregate species groups. Biomass estimates were calculated as the mean for the time period as:

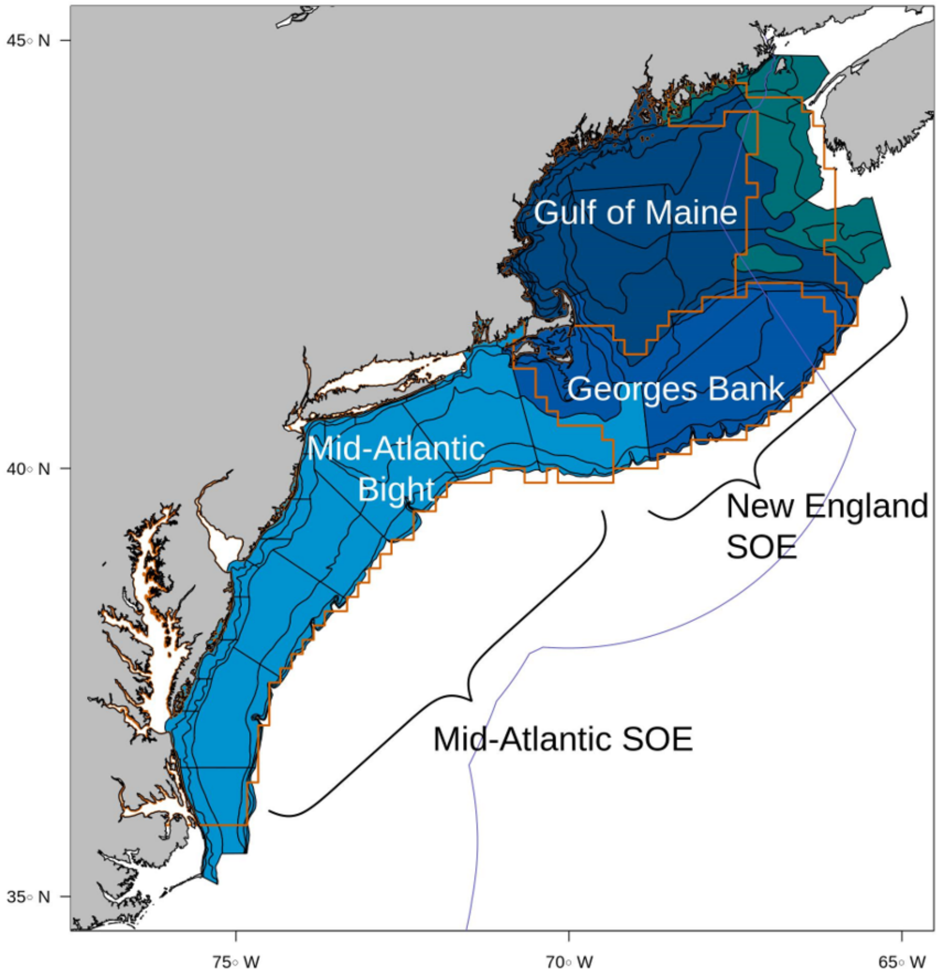


**Figure 2.** Map of the Northeast US Large Marine Ecosystem broken into Ecological Production Units (red line). The NEFSC bottom trawl survey strata are outlined in grey. The colors represent the strata used for the NEFSC State of the Ecosystem Report.

$$B_i = \frac{I}{q} \cdot \frac{A}{a} \tag{3}$$

where $I$ is the mean biomass per tow (converted to metric tons from kg), $A$ the area of the Georges Bank EPU, $a$ the average swept area of the tows, and $q$ the catchability. Data on catchability is sparse so $q$'s from EMAX (Link et al. 2006) were applied to scale biomass values. Swept area biomass was then converted to a density by dividing by the area of the Georges Bank EPU to be in the proper units for the mass balance model ($t\ km^{-2}$).

### 2.2.2. *Fisheries dependent data*

For the same time period as the fishery independent data, fisheries dependent data (landings and discards) were also obtained from the NEFSC. Landings were queried

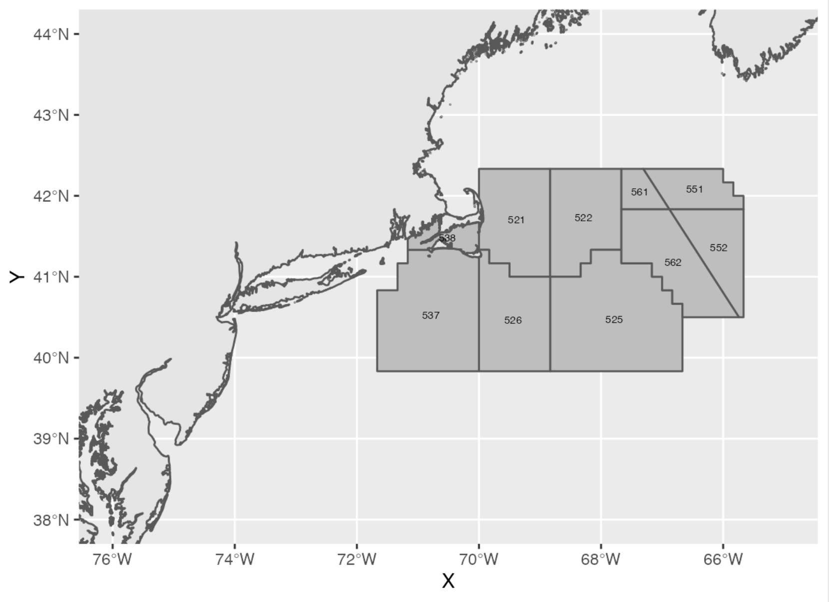


**Figure 3.** Map of the Northwest Atlantic Fisheries Organizations Statistical Areas used to define Georges Bank.

from the commercial fishery databases while discard ratios were calculated using observer data. Due to the nature of how data is collected there is a spatial mismatch between fishery independent and fisheries dependent data (Figure 3). Commercial landings are reported by Northwest Atlantic Fisheries Organization's Statistical Areas. For the purpose of this study we selected statistical areas 521, 522, 523, 524, 526, 551, 552, 561, and 562. As outlined in the MSKeyrun documentation, landings were proportioned between those inside and outside of the Georges Bank spatial footprint using the average ratio from the NEFSC bottom trawl survey.

Landings data were queried by species and gear (Table 2). Skates are reported as a mix. Fortunately most skates retained for human consumption are the larger Winter skate, *Leucoraja ocellata*, while the smaller Little skate, *Leucoraja erinacea*, is used as bait (NEFSC 2007). Therefore landings identified as *Skates Uncl* were segregated based on their utilization code. In addition to the skate reclassification, otter trawl landings were divided into small and large mesh fleets. These two similar gear types catch vastly different suites of species. Distinction between small and large was made using a mesh size of 6in (GARFO 2019). Landings were converted to metric tons and averaged over the time period. Landings were then divided by the area of the Georges Bank statistical areas to be in the proper units for the mass balance model ($t\ km^{-2}$).

Table 2.: Non-living groups for the Georges Bank Rpath model. Fleets were parameterized using Northeast Fishery Science Center (NEFSC) databases.

| Group | Type |
|---|---|
| Detritus | Detritus |
| Discards | Detritus |
| DredgeScallop | Fleet |
| DredgeClam | Fleet |
| Gillnet | Fleet |
| Longline | Fleet |
| Seine | Fleet |

Table 2.: Non-living groups for the Georges Bank Rpath model. Fleets were parameterized using Northeast Fishery Science Center (NEFSC) databases. *(continued)*

| Group | Type |
|---|---|
| PotTrap | Fleet |
| OttertrawlSmall | Fleet |
| OttertrawlLarge | Fleet |
| Midwater | Fleet |
| OtherFisheries | Fleet |

Observers record the disposition of catch aboard fishing vessels. The ratio of the discarded amount of a species to the amount kept from observed trips can be expanded by the total landings to estimate total discards as:

$$D_{ig} = \overline{\frac{\sum D_{ig}}{\sum K_g}} \cdot \sum L_g \tag{4}$$

where $D_{ig}$ are the discards for species $i$ and gear $g$, $K_g$ the retained or kept species, and $L_g$ the gear specific landings. Observers also record incidental takes of marine mammals and other protected species. Only records where a mortality was recorded were included as discards. Most of the incidental take records did not include a weight. Therefore an average weight for the species was used based on Trites and Pauly (1998). Similar to landings, discards were divided by the area of the Georges Bank statistical areas to be in the proper units for the mass balance model ($t\ km^{-2}$).

### *2.2.3. Food Habits*

The NEFSC has been collecting food habits data since 1973 (Link and Almeida 2000). There are over 1400 different prey items in the database so the first step was assigning all prey items to their respective species group. Next the food habits data was queried excluding blown and empty stomachs. Stomach contents identified as fish unclassified or animal remains were also ignored. Stomachs were used from the entire time series to ensure enough coverage of less well studied species groups.

Percent weight of prey was calculated using the cluster design explained in Nelson (2014) as:

$$\hat{r} = \frac{\sum_{i=1}^{n} M_i \hat{\mu_i}}{\sum_{i=1}^{n} M_i} \tag{5}$$

where $\hat{r}$ is the mean attribute of interest (in our case weight per stomach), $M_i$ is the total number of fish in each cluster (Station/Species group), and $\hat{\mu_i}$ the mean attribute in the cluster calculated as:

$$\hat{\mu_i} = \frac{\sum_{j=1}^{M_i} y_{ij}}{M_i} \tag{6}$$

where $y_{ij}$ is attribute of fish $j$ in cluster $i$. After calculating the mean weight per stomach of each prey item, percent diet by weight was calculated as:

$$\%prey_i = \frac{\hat{r_i}}{\sum_{i=1}^{n} \hat{r_i}} \tag{7}$$

### 2.2.4. *Upper and lower trophic levels*

Several species groups in the model are not well represented in the NEFSC surveys as outlined above. Most of these are upper trophic level species such as birds, sharks, and marine mammals or lower trophic levels like zooplankton and benthos (Table 3). For those groups we utilized the parameters in the EMAX model for Georges Bank (Link et al. 2006).

Table 3.: Living groups for the Georges Bank Rpath model that were parameterized using values from EMAX. Classifications and diet are those used for PREBAL calculations.

| Group | Classification | Diet |
|---|---|---|
| Seabirds | Bird | |
| Seals | Mammal | |
| BalWhale | Whale | |
| ToothWhale | Whale | |
| HMS | HMS | Piscivore |
| Sharks | Shark | Piscivore |
| Macrobenthos | Invertebrate (Benthic) | |
| Megabenthos | Invertebrate (Benthic) | |
| OtherShrimps | Invertebrate (Pelagic) | |
| Krill | Invertebrate (Pelagic) | |
| Micronekton | Invertebrate (Pelagic) | |
| GelZooplankton | Invertebrate (Pelagic) | |
| Mesozooplankton | Zooplankton | |
| Microzooplankton | Zooplankton | |
| Phytoplankton | Primary Producer | |

### 2.2.5. *Biological parameters*

Mass balance models require biological parameters that relate biomass to production and consumption. While more in-depth methods exist (see Aydin et al. 2007), for this model we relied on the Northwest Atlantic Continental Shelf (NWACS) Ecosystem model (Buchheister et al. 2017). This model was a shelf-wide model that used EMAX as a starting point but included many of the more disaggregated groups present in our model. The biological parameters used in Buchheister et al. (2017) were based on recent individual stock assessments so there was no need to duplicate effort. The NWACS model did have several multi-stanza groups whose parameter values were biomass weighted to be included. For groups not present in NWACS, we queried FishBase.org (FishBase 2019). Production to biomass was calculated a 1 over longevity and consumption to biomass was an average of the published studies on the site. After utilizing Buchheister et al. (2017) and FishBase.org, the aggregate groups *SmFlatfishes* and *OtherFlatfish* still did not have parameters. Due to their size and position in the food web, we used the *SmPelagics* for the *SmFlatfishes* and *OtherDemersals* for the *OtherFlatfish*.

### 2.3. *Balancing procedures*

Inevitably due to the uncertainty around parameter estimates models are not balanced. Within the mass balance framework, ecotrophic efficiencies ($EE$) represent the proportion of mortality explained in the model. Therefore values greater than one are considered out of balance. While getting a model to have no $EE$s greater than one is a good starting point, Link (2010) suggests some other criteria that should be explored. The pre-balance diagnostics (PREBAL) of Link (2010) are broken into five categories. These are: biomass across trophic levels, biomass ratios, vital rates across trophic levels, vital rate ratios, and total production and removals. These diagnostics follow general ecological and fishery principles.

The Georges Bank Rpath model was balanced by hand moving sequentially down from the largest $EE$ until the model was balanced. During the balancing process, PREBAL diagnostics were consulted. There are several ways to reduce the $EE$ of a species group. First you can increase the biomass or production of the group, second you can decrease the consumption on a group, or finally you can decrease the fishing pressure. Consumption can be lowered by reducing the biomass of the predator, its consumption to biomass ratio, or the proportion of the prey in its diet. Due to the large uncertainty around fishery independent data and the use of EMAX $q$s designed for different gear and more aggregate species, biomass was typically the first parameter to be manipulated. Production and consumption rates were held within range of similar continental shelf mass balance models obtained from the Ecobase website (Colléter et al. 2015). Finally, diet compositions were modified to alleviate predation.

Table 4.: Balanced Georges Bank model. The column type is the proportion of primary production used by the group (1 = primary producers, 0 = heterotroph) or 2 for detrital groups or 3 for fleets. The columns biomass and removals correspond to the mass per $km^2$. Other abbreviations are TL = Trophic Level; PB = Production to Biomass ratio; QB = Consumption to Biomass ratio; EE = Ecotrophic Efficiency (% mortality explained in the model); and GE = Growth Efficiency or Production to Consumption ratio.

| Group | type | TL | Biomass | PB | QB | EE | GE | Removals |
|---|---|---|---|---|---|---|---|---|
| Seabirds | 0 | 4.27 | 0.06 | 0.28 | 80.00 | 0.16 | 0.00 | 0.00 |
| Seals | 0 | 4.76 | 0.14 | 0.07 | 0.62 | 0.00 | 0.12 | 0.00 |
| BalWhale | 0 | 4.08 | 1.66 | 0.04 | 0.36 | 0.01 | 0.11 | 0.00 |
| ToothWhale | 0 | 4.85 | 0.49 | 0.04 | 1.59 | 0.09 | 0.03 | 0.00 |
| HMS | 0 | 4.27 | 0.14 | 0.58 | 2.26 | 0.01 | 0.26 | 0.00 |
| Sharks | 0 | 4.74 | 0.10 | 0.14 | 0.88 | 0.38 | 0.16 | 0.00 |
| AtlHerring | 0 | 3.79 | 6.32 | 1.10 | 3.70 | 0.87 | 0.30 | 0.02 |
| AtlMackerel | 0 | 3.98 | 0.39 | 0.55 | 2.17 | 0.96 | 0.25 | 0.02 |
| RiverHerring | 0 | 3.82 | 0.06 | 1.30 | 4.40 | 0.81 | 0.30 | 0.00 |
| Butterfish | 0 | 3.58 | 6.13 | 1.31 | 4.23 | 0.97 | 0.31 | 0.03 |
| SmPelagics | 0 | 3.22 | 11.24 | 1.64 | 5.47 | 0.80 | 0.30 | 0.00 |
| Mesopelagics | 0 | 3.36 | 0.13 | 1.10 | 3.70 | 0.80 | 0.30 | 0.00 |
| OtherPelagics | 0 | 4.65 | 1.34 | 0.59 | 2.19 | 0.80 | 0.27 | 0.00 |
| Cod | 0 | 4.36 | 6.68 | 0.11 | 0.37 | 0.97 | 0.29 | 0.40 |
| Haddock | 0 | 3.78 | 8.96 | 0.13 | 0.60 | 0.99 | 0.21 | 0.12 |
| Goosefish | 0 | 4.93 | 0.85 | 0.08 | 0.81 | 0.91 | 0.10 | 0.03 |
| OffHake | 0 | 4.83 | 0.03 | 0.45 | 2.44 | 0.75 | 0.18 | 0.00 |
| SilverHake | 0 | 4.51 | 14.35 | 0.24 | 0.76 | 0.96 | 0.32 | 0.10 |
| RedHake | 0 | 4.03 | 2.05 | 0.45 | 1.61 | 0.94 | 0.28 | 0.07 |
| WhiteHake | 0 | 5.05 | 0.57 | 0.11 | 0.62 | 0.70 | 0.18 | 0.01 |
| Redfish | 0 | 3.76 | 0.08 | 0.06 | 0.23 | 0.95 | 0.27 | 0.00 |

Table 4.: Balanced Georges Bank model. The column type is the proportion of primary production used by the group (1 = primary producers, 0 = heterotroph) or 2 for detrital groups or 3 for fleets. The columns biomass and removals correspond to the mass per $km^2$. Other abbreviations are TL = Trophic Level; PB = Production to Biomass ratio; QB = Consumption to Biomass ratio; EE = Ecotrophic Efficiency (% mortality explained in the model); and GE = Growth Efficiency or Production to Consumption ratio. *(continued)*

| Group | type | TL | Biomass | PB | QB | EE | GE | Removals |
|---|---|---|---|---|---|---|---|---|
| Pollock | 0 | 4.29 | 4.10 | 0.10 | 0.48 | 0.21 | 0.21 | 0.05 |
| OceanPout | 0 | 3.52 | 0.75 | 0.19 | 0.92 | 0.94 | 0.21 | 0.01 |
| BlackSeaBass | 0 | 3.98 | 0.03 | 0.12 | 0.83 | 0.30 | 0.15 | 0.00 |
| Bluefish | 0 | 4.84 | 0.65 | 0.28 | 1.42 | 0.44 | 0.20 | 0.00 |
| Scup | 0 | 3.40 | 0.26 | 0.45 | 1.61 | 0.82 | 0.28 | 0.01 |
| OtherDemersals | 0 | 3.74 | 3.26 | 0.52 | 1.74 | 0.90 | 0.30 | 0.02 |
| SouthernDemersals | 0 | 3.74 | 0.00 | 1.12 | 4.50 | 0.20 | 0.25 | 0.00 |
| Fourspot | 0 | 4.40 | 0.31 | 0.45 | 1.61 | 0.55 | 0.28 | 0.02 |
| SummerFlounder | 0 | 4.80 | 0.10 | 0.28 | 1.11 | 0.72 | 0.25 | 0.02 |
| AmPlaice | 0 | 3.69 | 0.27 | 0.12 | 0.92 | 0.90 | 0.14 | 0.02 |
| Windowpane | 0 | 3.93 | 3.71 | 0.14 | 0.92 | 0.99 | 0.16 | 0.02 |
| WinterFlounder | 0 | 3.47 | 3.55 | 0.18 | 0.92 | 0.90 | 0.19 | 0.10 |
| WitchFlounder | 0 | 3.39 | 0.08 | 0.12 | 0.92 | 0.95 | 0.14 | 0.01 |
| YTFlounder | 0 | 3.58 | 0.63 | 0.92 | 3.23 | 0.35 | 0.28 | 0.11 |
| OtherFlatfish | 0 | 3.74 | 0.25 | 0.52 | 1.74 | 0.80 | 0.30 | 0.00 |
| SmFlatfishes | 0 | 3.44 | 0.05 | 1.64 | 5.47 | 0.80 | 0.30 | 0.00 |
| SpinyDogfish | 0 | 4.42 | 28.24 | 0.08 | 0.29 | 0.29 | 0.28 | 0.12 |
| SmoothDogfish | 0 | 4.03 | 0.06 | 0.45 | 2.44 | 0.13 | 0.18 | 0.00 |
| Barndoor | 0 | 4.37 | 0.08 | 0.09 | 0.83 | 0.99 | 0.11 | 0.01 |
| WinterSkate | 0 | 4.06 | 18.43 | 0.12 | 0.42 | 0.59 | 0.29 | 0.17 |
| LittleSkate | 0 | 3.77 | 7.04 | 0.08 | 0.83 | 0.44 | 0.10 | 0.15 |
| OtherSkates | 0 | 3.87 | 1.23 | 0.10 | 0.42 | 0.80 | 0.24 | 0.00 |
| Illex | 0 | 3.86 | 0.44 | 5.72 | 19.00 | 0.67 | 0.30 | 0.01 |
| Loligo | 0 | 3.86 | 2.44 | 5.72 | 19.00 | 0.49 | 0.30 | 0.05 |
| OtherCephalopods | 0 | 3.86 | 0.17 | 5.72 | 19.00 | 0.80 | 0.30 | 0.00 |
| AmLobster | 0 | 3.02 | 2.60 | 0.17 | 1.29 | 1.00 | 0.14 | 0.02 |
| Macrobenthos | 0 | 2.36 | 138.67 | 2.40 | 16.84 | 0.82 | 0.14 | 0.00 |
| Megabenthos | 0 | 3.02 | 5.46 | 2.30 | 15.53 | 0.80 | 0.15 | 0.01 |
| AtlScallop | 0 | 2.08 | 8.09 | 1.20 | 6.66 | 0.41 | 0.18 | 0.55 |
| Clams | 0 | 2.08 | 55.22 | 1.20 | 6.66 | 0.32 | 0.18 | 0.04 |
| OtherShrimps | 0 | 2.63 | 4.76 | 2.00 | 6.66 | 0.80 | 0.30 | 0.00 |
| Krill | 0 | 2.92 | 12.00 | 14.25 | 85.50 | 0.38 | 0.17 | 0.00 |
| Micronekton | 0 | 2.92 | 18.40 | 14.25 | 85.50 | 0.53 | 0.17 | 0.00 |
| GelZooplankton | 0 | 3.09 | 20.96 | 20.00 | 100.00 | 0.72 | 0.20 | 0.00 |
| Mesozooplankton | 0 | 2.32 | 121.20 | 46.00 | 108.97 | 0.92 | 0.42 | 0.00 |
| Microzooplankton | 0 | 2.32 | 24.80 | 85.00 | 212.55 | 0.77 | 0.40 | 0.00 |
| Bacteria | 0 | 2.00 | 20.74 | 91.25 | 285.16 | 0.79 | 0.32 | 0.00 |
| Phytoplankton | 1 | 1.00 | 79.09 | 174.35 | 0.00 | 0.82 | 0.00 | 0.00 |
| Detritus | 2 | 1.00 | 29404.66 | 0.50 | 0.00 | 0.80 | 0.00 | 0.00 |
| Discards | 2 | 1.00 | 1.64 | 0.50 | 0.00 | 0.05 | 0.00 | 0.00 |
| ScallopDredge | 3 | 3.27 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| ClamDredge | 3 | 3.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| OtherDredge | 3 | 3.18 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| FixedGear | 3 | 5.29 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Pelagic | 3 | 4.78 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Trap | 3 | 4.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| SmallMesh | 3 | 5.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| LargeMesh | 3 | 4.99 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| HMSFleet | 3 | 5.37 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

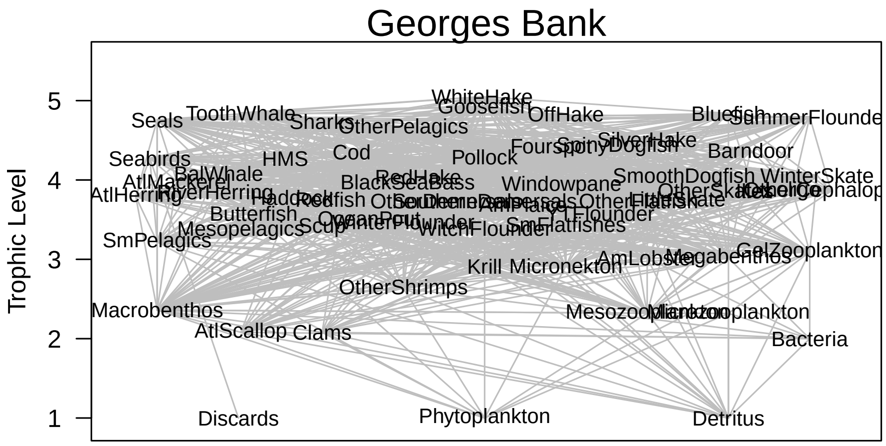


**Figure 4.** Food web of the Georges Bank Rpath model.

Table 4.: Balanced Georges Bank model. The column type is the proportion of primary production used by the group (1 = primary producers, 0 = heterotroph) or 2 for detrital groups or 3 for fleets. The columns biomass and removals correspond to the mass per $km^2$. Other abbreviations are TL = Trophic Level; PB = Production to Biomass ratio; QB = Consumption to Biomass ratio; EE = Ecotrophic Efficiency (% mortality explained in the model); and GE = Growth Efficiency or Production to Consumption ratio. *(continued)*

| Group | type | TL | Biomass | PB | QB | EE | GE | Removals |
|---|---|---|---|---|---|---|---|---|
| OtherFisheries | 3 | 5.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

# 3. Results

## 3.1. *Georges Bank model structure*

The Georges Bank Rpath model consists of 71 groups (Tables 4). Of these, 59 are living groups comprised of individual fish and invertebrate species, aggregate fish and invertebrate groups, marine mammals, birds, primary and secondary producers, and bacteria. There are two detrital groups representing discards and general detritus. Finally, there are 10 fleets representing the various fishing gears used on Georges Bank. The resultant food web (Figure 4) is highly interconnect, an expected result due to the generalist nature of many of the species.

## 3.2. *Balancing*

PreBal provides some rules of thumb that put your model in a realistic starting point (Link 2010). They are also helpful during the balancing procedure. The first criteria for

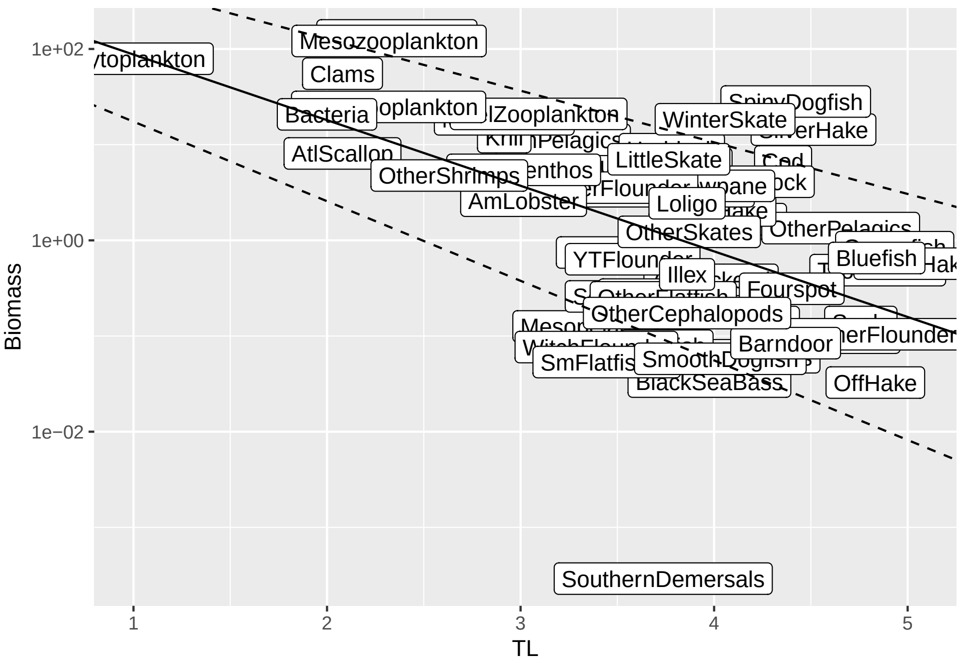


**Figure 5.** Biomass decomposition over trophic level. Biomass spans six orders of magnitude and has a slope of -0.69.

PREBAL is biomass across trophic levels. The initial biomass span was six orders of magnitude which is in line with PreBal suggestion of five to seven. The decomposition slope was 69% which is off from the 5-10% suggestion of PreBal (Figure 5). The reason is that PreBal uses a non-parametric rank for the x-axis whereas I used the species trophic level. This value makes much more sense for a system that spans six orders of magnitude from phytoplankton to top predators. This should be investigated more. Biomasses for some fringe species (well below the line on figure 5) were not increase purposely. All of these groups represent species that are only partially on the bank. The bulk of their biomass resides outside the bank.

During the balancing procedure, the largest deficiencies in *EE* were addressed first. Many of these groups were poorly sampled or aggregated (or both). This included several pelagic groups who are not sampled well in the survey. To simplify issues that arose from aggregate groups with little to no data available their EEs were set to 0.8. Undersampled pelagic species were increased by an order of magnitude. In addition, all species biomasses were multiplied by 4 to better scale with fisheries data.

Several key predators needed to have their biomass reduced as well. The most significant of these was *SpinyDogfish*. This was a similar issue in Buchheister et al. (2017) most likely arising from an aggregate $q$ that is not really applicable to their behavior with a trawl. There were a few situations were diet compositions were adjusted to help achieve balance. Once again this mostly occurred with the aggregated species groups such as *OtherFlatfish* and *OtherCephalopods*. In these cases a portion of the diet composition assigned to the species were moved to a similar species. The other notable diet composition change occurred for three hake species groups (*RedHake*, *SilverHake*, and *WhiteHake*) where conspecific mortality was reduced. Cannibalism can be hard to balance in these models as increasing biomass actually has an adverse effect. So species with lower productivities (fish species) have a hard time sustaining more than

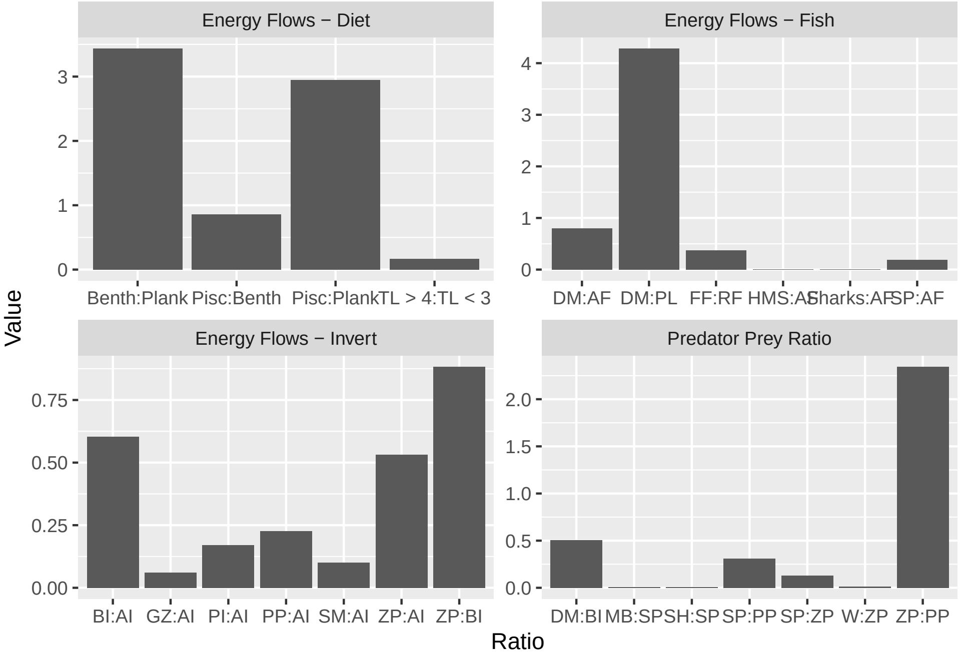


**Figure 6.** Biomass ratios for the Georges Bank Rpath model. Top left panel displays ratios between predators and their prey. Top right panel and bottom left display the relative energy flow through the system as ratios of aggregate groups to total fish (top right) or invertebrates (bottom left). The bottom right panel displays the ratio between fish trophic guilds. Abbreviations are: AF - all fish, AI - all invertebrates, Benth - benthivores, BI - benthic invertebrates, DM - demersal fish, FF - flatfish, GZ - gelatinous zooplankton, HMS - highly migratory species, MB - mammals and birds, PI - pelagic invertebrates, Pisc - piscivores, PL - pelagics, Plank - planktivores, PP - primary producers, SH - sharks and HMS, SM - shrimp, micronekton, and krill, SP - small pelagics, TL - trophic level, W - whales, ZP - zooplankton.

10%.

The second criteria looks at ratios between taxa groups (Figure 6). Initial impressions led to an increase in *Phytoplankton* which improved the ratios to primary production. Several of the ratios are greater than one which is indicative of too much predation on a prey group (Link 2010). However, the generalist nature of many of the predators in the system means that not one prey group is experiencing all of the predation pressure from a particular group of predators. Very small ratios are indicative of multiple trophic levels between the predator and the prey and are expected for groups like whales:zooplankton (Figure 6:A).

The third and fourth criteria are similar to criteria one and two but for vital rates (consumption, production, respiration) instead of biomass. Similar to the biomass plots, the slopes were steeper than suggested. Once again this is most likely due to the high productivity of Georges Bank. There was no attempt made to correct this issue as it would have meant dramatically lowering the vital rates of lower trophic levels or raising the vital rates of mid to upper trophic levels. Most vital rates were within range of other studies obtained from Ecobase. Vital rate decomposition plots were consulted during the balancing process to identify outliers (Not shown).

The fifth criteria of the PREBAL diagnostics evaluates overall production and consumption. The model did not have any groups with growth efficiencies (GE: Table 4 or production to consumption) that were too low or high or any with ratios greater than one. This is good as it is not physically possible to have more production than consumption (Link 2010).

## 4. Discussion

Georges Bank is a highly productive region of the world's ocean; highlighted by the diverse fauna present throughout the bank and long storied history of fishing. Our mass balance representation of Georges Bank captures that essence. Based on the PREBAL rules of thumb for balancing marine ecosystems, Georges Bank has a steeper biomass and vital rate decomposition than typical marine systems (Link 2010). This is indicative of the system having high primary production as well as lower biomass and production from the mid to upper trophic levels. In addition, the food web is highly interconnected. This aligns with the fact that most of the predators on Georges Bank are generalists, eating whatever prey is available (Link and Almeida 2000).

The distinct geographic and oceanographic properties of Georges Bank make it an ideal area to enact EBFM. Of course there will be many questions surrounding the implementation of place-based strategies. Some of which could be addressed through an MSE process. Our Georges Bank Rpath model could serve as an operating model or contribute to a suite of operating models for that purpose. Rpath has been designed with these types of applications in mind (Lucey et al. 2021).

As a full ecosystem model, mass balance models can encapsulate most of the species in a system. Although with any modeling approach it is not necessary to have explicit details for every trophic level (Fulton, Smith, and Johnson 2003). The computational overhead for Rpath is fairly low which allows mass balance models to be included in sensitivity analysis (Gaichas, Aydin, and Francis 2015). The compromise is losing the ability to explore length (or age) and/or spatial questions. While not part of this study, mass balance models can be parameterized to include multi-stanza groups which could mimic the length stanzas in Hydra (Lucey, Gaichas, and Aydin 2020). Mass balance models can also be extended to include spatial properties if using the Ecospace algorithms of the Ecopath with Ecosim software (Christensen and Walters 2004).

There have been a number of mass balance models produced for the Northeast US Continental Shelf Large Marine Ecosystem. Unfortunately they are too aggregated (i.e. Link et al. 2006) or not specific to the Georges Bank region (i.e. Buchheister et al. 2017). Therefore we used the Link et al. (2006) work as a starting point supplemented by Buchheister et al. (2017). The new fleet structure and species disaggregation will allow for more management policies to be explored.

Our Georges Bank model is the static snapshot of energy flow through the system. For management purposes it should be tuned to data and ideally designated as a 'key-run' (ICES 2016). 'Key-runs' serve as a quality control device that provide confidence to managers that the model is appropriate for use in providing advice. The next step in creating a 'key-run' will be to tune the model to data.

This Georges Bank Rpath model is a good starting point for management. The complexity of the model allows it to be flexible and inclusive. Inevitably during the MSE process specific questions will need to be addressed. At that time further refinement of the model and its parameters can be made. This model refinement should take place as part of a stakeholder process to ensure the model is capturing the aspects of the system in which they are interested (Fulton et al. 2014; Goethel et al. 2018).

## References

Aydin, KY, S. Gaichas, I Ortiz, D Kinzey, and N Friday. 2007. *A comparison of the Bering Sea, Gulf of Alaska, and Aleutian Islands large marine ecosystems through food web modeling.* US Department of Commerce, NOAA Technical Memo NMFS-AFSC-178.

Buchheister, Andre, Thomas J Miller, Edward D Houde, and David A Loewensteiner. 2017. *Technical Documentation of the Northwest Atlantic Continental Shelf (NWACS) Ecosystem Model. Report to the Lenfest Ocean Program, Washington, D.C.* University of Maryland Center for Environmental Sciences Report TS -694 -17.

Christensen, Villy, and Daniel Pauly. 1992. "ECOPATH II - a software for balancing steady-state ecosystem models and calculating network characteristics." *Ecological Modelling* 61: 169–185.

Christensen, Villy, and Carl J Walters. 2004. "Ecopath with Ecosim: methods, capabilities and limitations." *Ecological Modelling* 172 (2–4): 109–139.

Colléter, Mathieu, Audrey Valls, Jérôme Guitton, Didier Gascuel, Daniel Pauly, and Villy Christensen. 2015. "Global overview of the applications of the Ecopath with Ecosim modeling approach using the EcoBase models repository." *Ecological Modelling* 302: 42–53. Accessed 2019-08-06. `http://www.sciencedirect.com/science/article/pii/S0304380015000447`.

FishBase. 2019. "FishBase." Accessed 2019-08-06. `https://www.fishbase.org`.

Fulton, E. A., A. D. M. Smith, and C. R. Johnson. 2003. "Effect of complexity on marine ecosystem models." *Marine Ecology-Progress Series* 253: 1–16. `://000183487200001`.

Fulton, Elizabeth A., Anthony D. M. Smith, David C. Smith, and Penelope Johnson. 2014. "An Integrated Approach Is Needed for Ecosystem Based Fisheries Management: Insights from Ecosystem-Level Management Strategy Evaluation." *PLOS ONE* 9 (1): e84242. Accessed 2019-06-28. `https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0084242`.

Gaichas, Sarah, Kerim Aydin, and Robert C. Francis. 2015. "Wasp waist or beer belly? Modeling food web structure and energetic control in Alaskan marine ecosystems, with implications for fishing and environmental forcing." *Progress in Oceanography* 138: 1–17.

GARFO. 2019. "Northeast (NE) multispecies small mesh fishery exemptions." Accessed 2019-08-06. `https://www.greateratlantic.fisheries.noaa.gov/regs/infodocs/small_mesh_exemption.pdf`.

Goethel, Daniel R., Sean M. Lucey, Aaron M. Berger, Sarah K. Gaichas, Melissa A. Karp, Patrick D. Lynch, John F. Walter, Jonathan J. Deroba, Shana Miller, and Michael J. Wilberg. 2018. "Closing the feedback loop: on stakeholder participation in management strategy evaluation." *Canadian Journal of Fisheries and Aquatic Sciences* 1–19. Accessed 2019-08-07. `http://www.nrcresearchpress.com/doi/10.1139/cjfas-2018-0162`.

ICES. 2016. *Report of the Working Group on Multispecies Assessment Methods (WGSAM), 9–13 November 2015, Woods Hole, USA.* ICES CM 2015/SSGEPI:20.

Jacobson, Larry, and Daniel Hennen. 2019. *Improving the NEFSC Clam Survey for Atlantic Surfclams and Ocean Quahogs.* Northeast Fisheries Science Center Reference Document 19-06.

Link, J. S., and F. P. Almeida. 2000. *An Overview and History of the Food Web Dynamics Program of the Northeast Fisheries Science Center, Woods Hole, Massachusetts.* NOAA Technical Memorandum NMFS-NE-159.

Link, Jason, Laurel Col, Vincent Guida, David Dow, John O'Reilly, Jack Green, William Overholtz, et al. 2009. "Response of balanced network models to large-scale perturbation: Implications for evaluating the role of small pelagics in the Gulf of Maine." *Ecological Modelling* 220 (3): 351–369. `https://www.sciencedirect.com/science/article/pii/S0304380008004973`.

Link, Jason S. 2010. "Adding rigor to ecological network models by evaluating a set of pre-balance diagnostics: A plea for PREBAL." *Ecological Modelling* 221 (12): 1580–1591.

Link, Jason S., Carolyn A. Griswold, Elizabeth T. Methratta, and Jessie Gunnard. 2006.

"Documentation for the Energy Modeling and Analysis eXercise (EMAX)." .

Lucey, SM, KY Aydin, SK Gaichas, SX Cadrin, G Fay, MJ Fogarty, and A Punt. 2021. "Evaluating fishery management strategies using an ecosystem model as an operating model." *Fisheries Research* 234.

Lucey, SM, SK Gaichas, and KY Aydin. 2020. "Conducting reproducible ecosystem modeling using the open source mass balance model Rpath." *Ecological Modelling* 427.

NEFSC. 2007. *44th SAW Assessment Summary Report.* Northeast Fisheries Science Center Reference Document 07-03.

NEFSC. 2018. *65th Northeast Regional Stock Assessment Workshop (65th SAW) Assessment Report.* Northeast Fisheries Science Center Reference Document 18-11.

Nelson, Gary A. 2014. "Cluster Sampling: A Pervasive, Yet Little Recognized Survey Design in Fisheries Research." *Transactions of the American Fisheries Society* 143 (4): 926–938.

Plaganyi, E. E. 2007. *Models for an ecosystem approach to fisheries.* FAO Fisheries Technical Paper 477. Rome: FAO.

Politis, Philip J., John K. Galbraith, Paul Kostovick, and Russell W. Brown. 2014. *Northeast Fisheries Science Center bottom trawl survey protocols for the NOAA Ship Henry B. Bigelow.* Northeast Fisheries Science Center Reference Document 14-06. Accessed 2019-08-06. `https://repository.library.noaa.gov/view/noaa/4825`.

Polovina, J. J. 1984. "Model of a Coral-Reef Ecosystem .1. The Ecopath Model and Its Application to French Frigate Shoals." *Coral Reefs* 3 (1): 1–11.

Trites, Andrew W, and Daniel Pauly. 1998. "Estimating mean body masses of marine mammals from maximum body lengths." 76: 11.

Ulanowicz, Robert E., and James J. Kay. 1991. "A package for the analysis of ecosystem flow networks." *Environmental Software* 6 (3): 131–142. `http://www.sciencedirect.com/science/article/pii/026698389190024K`.

Walters, C., V. Christensen, and D. Pauly. 1997. "Structuring dynamic models of exploited ecosystems from trophic mass-balance assessments." *Reviews in Fish Biology and Fisheries* 7 (2): 139–172.

Walters, C., D. Pauly, and V. Christensen. 2000. "Ecospace: prediction of mesoscale spatial patterns in trophic relationships of exploited ecosystems, with emphasis on the impacts of marine protected areas." *Ecosystems* 2: 539—554.