**North Sea herring benchmark**

**Preliminary investigation of the effect of changes in the input data and model configuration**

**Benoit Berges, Thomas Brunel, Niels Hintzen,**

**Wageningen Marine Research**

Contents

[1 Introduction 2](#_Toc505350757)

[2 Methods 3](#_Toc505350758)

[3 Assessment update 4](#_Toc505350759)

[3.1 Update of the proportion of fishing mortality occurring before spawning 4](#_Toc505350760)

[3.2 Update of the natural mortality matrix 6](#_Toc505350761)

[3.3 Update of IBTS-Q1 index time series at age 1: 10](#_Toc505350762)

[3.4 Use of IBTS-Q1 index time series at age 1 to 5: 14](#_Toc505350763)

[3.5 Including IBTS-Q3 index time series at all ages 18](#_Toc505350764)

# Introduction

In preparation of the benchmark assessment for the North Sea Autumn Spawning herring, WMR investigated the effect of a number of changes in the input data and model configuration on the assessment :

1. update the proportion of fishing mortality occurring before spawning (from a constant value to annual values).
2. Update of the natural mortality time series based on the most recent (2016) run of SMS
3. Update of the IBTS index (new estimation method) and incorporation of additional age-classes in the assessment :
   1. Updating to new IBTS\_Q1 series for age 1
   2. Adding older ages for IBTS\_Q1
   3. Adding IBTS\_Q3 for all ages, leaving aggregation with default settings
   4. (Optional) changing aggregation settings
4. Test the new model version in which the proportion of each spawning component is explicitly modelled and the SSB index is expressed per component.

# Methods

The impact of each change made on the assessment (data used or model configuration) was evaluated by a comparison of a series of diagnostics between the updated model and the base case. A sequential approach was adopted in which changes in the assessment are accepted (or rejected) successively, and once a change is accepted, it is used as the new base case for evaluation of further changes.

The initial base case used is the 2017 assessment.

The diagnostics used to compare models were :

1. model parameters: they inform on how the model treats the observations (parameter observation variance), on the modelled processes (process error and random walks variances) and on the scaling between modelled abundances and survey indices (catchabilities)
2. model likelihood (goodness of fit), which was used to test the significance of additional parameters in nested models (fitted on the same data) when different model configurations are tested
3. residuals for the observations : patterns (temporal or along ages) indicate departure from the assumption of an independent identical distribution. Any changes in the assessment that would impact the residuals should be scrutinised.
4. model uncertainty : changes in the assessment can affect the quality of the fit, and therefore the precision with which parameters are estimated. Standard deviation (or here confidence intervals) of the parameter estimates can be compared between 2 assessments. The strength of parameter correlations can also be compared between 2 models. uncertainty in parameters reflect in uncertainty in the states, we are investigated here using the CV of essential quantities defining stock status (SSB, Fbar and recruitment).
5. model stability : comparison of retrospective plot between two assessments will highlight any change in model stability

All the R scripts, model output and figures are available on Github (<https://github.com/ices-eg/wg_HAWG/tree/master/NSAS/benchmark>).

# Assessment update

## Update of the proportion of fishing mortality occurring before spawning

differences

The proportion of fishing mortality occurring before spawning (Fprop) time is used in the model to compute the number of individuals at spawning time, and therefore SSB. This data has been updated from a single value constant through time, to a time varying vector (but remains age-invariant). Figure 3.1‑1 shows a comparison of the Fprop vector used in the HAWG2017 assessment (ref) with the new created vector. The new values fluctuate around 0.62, just under the value of the previous assessment, i.e. 0.67.

interpretation

Given the limited magnitude of the changes on the input vector for Fprop, and the fact that the fit to the only SSB index (SCAI) is very poor in the model, the SAM model fitted on this new Fprop data (named 1\_newFprop) is almost identical to the last HAWG assessment (Figure 3.1‑2). There is only minor differences (and certainly none of significance) in estimated model parameters (not shown).

Conclusion :

The revision of the Fprop data has virtually no effect on the assessment, and since the new data is an improvement on the constant value used in the past, this change in the assessment should be accepted. In case of significant change in activity of the fishery in the future, or in the time of spawning, it is important to incorporate this information in the assessment. It is therefore preferable to update each year the values of Fprop, instead of using a constant value.

|  |
| --- |
|  |
| **Figure 3.1‑1:** revision of the proportion of fishing mortality occurring before spawning (0\_basecase = 2017 assessment at HAWG, 1\_newFprop = revised data). |

|  |
| --- |
|  |
| **Figure 3.1‑2:** comparison of the stock trajectories for the assessment using the old and the new Fprop data. |

## Update of the natural mortality matrix

differences

The natural mortality at age matrix was updated on the basis of a new run of the SMS model carried out in 2017. The method used to compute the North Sea herring mortality at age matrix from the output of SMS was the same as applied for the last update. The updated mortality data show lower values for most age-classes throughout the time series (Figure 3.2‑1). The only exception is for age 0, for which the new mortality estimate becomes larger than the old one after 1985.

The SAM model was fitted using the new maturity data (named 2\_newM) and compared to the previous model (1\_newFprop). The changes in the natural mortality matrix did not affect drastically the model characteristics. Parameter estimates are very similar (none is significantly different between model 1\_newFprop and 2\_newM). The largest differences are found for the catchabilities of the surveys which are all revised upwards (Figure 3.2‑2). Observation and process variances are very similar, except for a slightly higher observation variance for the catches and lower random walk variance for F (smoother variations) for the model with the new mortality data (Figure 3.2‑3).

The residuals to all sources of observations were identical and the retrospective analysis did not show any noticeable difference (not shown).

There was no marked differences in the uncertainty on the parameter estimates (Figure 3.2‑2 and Figure 3.2‑3). The overall model uncertainty on the quantities relevant to management (SSB, Fbar, Figure 3.2‑4) was slightly reduced with the new mortality data, but the difference became more substantial for the early years in the time series (prior to 1965).

Finally, using the revised natural mortality resulted in a revised perception of the historic stock development, with a SSB around 16% lower for the recent decades and up to 25% for the early part of the time series (Figure 3.2‑5). Conversely the Fbar values are estimated around 20% higher for the recent decades and up to 30% for the older years.

Interpretation

Catch and survey observations in the model essentially give information on the total mortality. Given that the new natural mortality is lower, the model tends to compensate it by a higher fishing mortality. This higher F can only produce the same catches if the underlying stock numbers are smaller, which explain the overall difference in SSB and Fbar trends observed on Figure 3.2‑5). This rescaling of the stock explains the differences in estimated survey catchabilities.

The cause of the changes in the catches observation variance and F random walk variance is more difficult to comprehend, however, those differences are small. We can just notice that these two paramters are inherently negatively correlated (a tighter fit to the catch data would imply a more variable F).

Conclusions

The changes in the natural mortality matrix reduced model uncertainty, but reduced the fit to the catch data (marginally in both cases) and it is therefore not possible to clearly say whether the model improved or deteriorated.

The new natural mortality data, according to ICES standards, represents the best available information (as being the latest data validated by ICES WGSAM) and should therefore be used for the North Sea herring assessment.

However it is important to point out that, given the sensitivity of the NSAS assessment to the revision of the SMS model shown here, the uncertainty in the output of the SMS model adds uncertainty on the NSAS assessment (in addition to the SAM model uncertainty).

|  |
| --- |
|  |
| Figure 3.2‑1: comparison of the previous and the updated natural mortality at age for North Sea herring. |

|  |
| --- |
| C:\Users\brune001\my git files\wg_HAWG\NSAS\benchmark\results\2_newM\comparison of catchabilities.png |
| Figure 3.2‑2: comparison of estimated survey catchability for the assessments with the old and with the updated natural mortality. |

|  |  |
| --- | --- |
| (a)  C:\Users\brune001\my git files\wg_HAWG\NSAS\benchmark\results\2_newM\comparison of obs.vars.png | (b)  C:\Users\brune001\my git files\wg_HAWG\NSAS\benchmark\results\2_newM\comparison of process.vars.png |
| Figure 3.2‑3: comparison of estimated observation and process variances for the assessments with the old and with the updated natural mortality | |

|  |
| --- |
| C:\Users\brune001\my git files\wg_HAWG\NSAS\benchmark\results\2_newM\comparison of model uncertainty.png |
| **Figure 3.2‑4:** comparison of uncertainty (CV) of Fbar, SSB and Recruitment for the assessments with the old and with the updated natural mortality |

|  |
| --- |
| C:\Users\brune001\my git files\wg_HAWG\NSAS\benchmark\results\2_newM\comparison of stock trajectories.png |
| Figure 3.2‑5: comparison of stock trajectories (Fbar, SSB and Recruitment) for the assessments with the old and with the updated natural mortality. |

## Update of IBTS-Q1 index time series at age 1:

Differences

The IBTS-Q1 index time series was updated during the North Sea Herring data workshop in December 2017 (ref). The two time series are shown in Figure 3.3‑1 for age 1. It can be observed that the revision of the index resulted in significant changes in scale. This is due to the new standardization of the index. To date, only age 1 was used in the assessment and this section will investigate the impact of the updated time series. This will be done with the results of the assessment resulting from Section 3.2 (update of the natural mortality index). In the various figures, the results from the previous section will be labelled as “2\_newM” while the results using the updated IBTS-Q1 time series will be labelled “3a\_newIBTSQ1\_age1”.

Interpretation

Figure 3.3‑2 shows that the inclusion of the new time series has only a minor impact on the variance of observed and unobserved variables. The change in uncertainty for Fbar is minor while it is worsened for recruitment and SSB, especially for years prior to 1990 (Figure 3.3‑3). The resulting stock trajectory (SSB, recruitment) exemplifies only minor differences (Figure 3.3‑4).

Conclusion

The changes introduced with the new IBTS-Q1 indices are substantial with a revised calculation method (ref). The impact on the overall assessment is minor and it is preferable to use the updated index as it is now standardized and more representative.

|  |
| --- |
| D:\git\wg_HAWG\NSAS\benchmark\results\3a_newIBTSQ1_age1\comparison of IBTSQ1 time series age 1 .png |
| Figure 3.3‑1: comparison of IBTS-Q1 age 1 time series. |

|  |  |
| --- | --- |
| (a)  D:\git\wg_HAWG\NSAS\benchmark\results\3a_newIBTSQ1_age1\newIBTSQ1_age1 figures_comparison - 03.png | (b)  D:\git\wg_HAWG\NSAS\benchmark\results\3a_newIBTSQ1_age1\newIBTSQ1_age1 figures_comparison - 04.png |
| Figure 3.3‑2: variance for observed and unobserved variables. | |

|  |
| --- |
| D:\git\wg_HAWG\NSAS\benchmark\results\3a_newIBTSQ1_age1\newIBTSQ1_age1 figures_comparison - 05.png |
| **Figure 3.3‑3:** assessment uncertainty |

|  |
| --- |
| D:\git\wg_HAWG\NSAS\benchmark\results\3a_newIBTSQ1_age1\newIBTSQ1_age1 figures_comparison - 01.png |
| Figure 3.3‑4: stock trajectory |

## Use of IBTS-Q1 index time series at age 1 to 5:

Differences:

To date, the assessment only used the IBTS-Q1 at age 1. Following the update of the age 1 time series (Section 3.3), this sections investigate the inclusion of age 2 to 6 into the assessment. The time series for each age (old and revisited) is shown in Figure 3.4‑1. It is important to note that there is no records available for age 6 for the old time series. Similarity in trends is variable. While strong for age 1, 2 and 3, the time series for ages 4 and 5 exemplify different trends, especially for recent years (2010 onward). The results from Section 3.3 will be labelled “3a\_newIBTSQ1\_age1” while the results using all the ages of the revised IBTSQ1 index will be labelled “3b\_newIBTSQ1\_allAges”. Worth including internal consistency of the data plot.

Interpretation:

The level of fitting of the SAM model is shown in Figure 3.4‑2 with a plot of the residuals for each age at every year. There seems to be correlations between the residuals between age 3 to 6 for early years. Of importance when including a new source of data is the observation variance as it determines the weight of this source in the final assessment. This is shown in Figure 3.4‑2. It can be observed that the newly added indices (age 2 to 6) exemplify the highest amount of variance which in turn will lower their impact on the final results of the assessment.

Conclusion

The addition of the new IBTSQ1 time series do no impact the assessment significantly. However, this is due to the high observation variance for this new data source. While these new data should be included in the assessment, it is important to investigate the source of this discrepancy.

|  |  |
| --- | --- |
| **D:\git\wg_HAWG\NSAS\benchmark\results\3b_newIBTSQ1_allAges\time series age -1.png** | **D:\git\wg_HAWG\NSAS\benchmark\results\3b_newIBTSQ1_allAges\time series age -2.png** |
| **D:\git\wg_HAWG\NSAS\benchmark\results\3b_newIBTSQ1_allAges\time series age -3.png** | **D:\git\wg_HAWG\NSAS\benchmark\results\3b_newIBTSQ1_allAges\time series age -4.png** |
| **D:\git\wg_HAWG\NSAS\benchmark\results\3b_newIBTSQ1_allAges\time series age -5.png** | **D:\git\wg_HAWG\NSAS\benchmark\results\3b_newIBTSQ1_allAges\time series age -6.png** |
| Figure 3.4‑1: IBTSQ1 time series. | |

|  |
| --- |
| D:\git\wg_HAWG\NSAS\benchmark\results\3b_newIBTSQ1_allAges\newIBTSQ1_allAges figures_comparison - 06.png |
| **Figure 3.4‑2:** residual by year |

|  |
| --- |
|  |
| **Figure 3.4‑3:** Observation variance by data source |

|  |
| --- |
| D:\git\wg_HAWG\NSAS\benchmark\results\3b_newIBTSQ1_allAges\newIBTSQ1_allAges figures_comparison - 05.png |
| Figure 3.4‑4: assessment uncertainty. |

|  |
| --- |
| D:\git\wg_HAWG\NSAS\benchmark\results\3b_newIBTSQ1_allAges\newIBTSQ1_allAges figures_comparison - 01.png |
| Figure 3.4‑5: stock trajectory for recruitment, SSB, catches and harvest |

## Including IBTS-Q3 index time series at all ages

Differences

In this section, the inclusion of a new survey data source, IBTS-Q3, is investigated. The time series for this new data source is shown in Figure 3.5‑1 together with other survey data available for the assessment (i.e. IBTS-Q1 and HERAS). In the various figures, the results from the previous section will be labelled as “3b\_newIBTSQ1\_allAges” while the results using the IBTS-Q3 time series will be labelled “3c\_IBTSQ3”.

Interpretation

From the time series shown in Figure 3.5-1, it can be observed that the time series from the different surveys have similar trends. The fitting from the SAM model of the time series is shown as the residuals for each age at every year in Figure 3.5-2. No clear correlation can be observed between the residuals. Further, the computed variance for each data source is plotted in Figure 3.5-3. The performance for the variance is varying depending on the ages. Ages 3 and 6 are the most informative while age 0 is found to be at the level of the SCAI index, having low impact in the assessment. The resulting uncertainty and stock trajectories are plotted in Figure 3.5-4 and Figure 3.5-5 respectively. Including the IBTS-Q3 results in a slight improvement in uncertainty (Figure 3.5-4) and the impact on the assessment is minor (Figure 3.5-5).

Conclusion

The index from the IBTS-Q3 survey provides further information for the stock assessment and show similar trends to other surveys (IBTS-Q1, HERAS). The overall impact of including this new data source is beneficial in term of uncertainty improvement and does not impact the assessment outputs significantely.

|  |  |
| --- | --- |
| **D:\git\wg_HAWG\NSAS\benchmark\results\3c_IBTSQ3\time series age - 1.png** | **D:\git\wg_HAWG\NSAS\benchmark\results\3c_IBTSQ3\time series age -2.png** |
| **D:\git\wg_HAWG\NSAS\benchmark\results\3c_IBTSQ3\time series age -3.png** | **D:\git\wg_HAWG\NSAS\benchmark\results\3c_IBTSQ3\time series age -4.png** |
| **D:\git\wg_HAWG\NSAS\benchmark\results\3c_IBTSQ3\time series age -5.png** | **D:\git\wg_HAWG\NSAS\benchmark\results\3c_IBTSQ3\time series age -6.png** |
| Figure 3.5‑1: comparison of survey time series across all ages .Age 6 for the HERAS survey represents the age plus group, i.e. the aggregation of age 6 to 9. | |

|  |
| --- |
| D:\git\wg_HAWG\NSAS\benchmark\results\3c_IBTSQ3\IBTSQ3 figures_comparison - 06.png |
| Figure 3.5‑2: residual by year. |

|  |
| --- |
| D:\git\wg_HAWG\NSAS\benchmark\results\3c_IBTSQ3\IBTSQ3 figures - 38.png |
| Figure 3.5‑3: observation variance by data source. |

|  |
| --- |
| D:\git\wg_HAWG\NSAS\benchmark\results\3c_IBTSQ3\IBTSQ3 figures_comparison - 05.png |
| Figure 3.5‑4: assessment uncertainty |

|  |
| --- |
| D:\git\wg_HAWG\NSAS\benchmark\results\3c_IBTSQ3\IBTSQ3 figures_comparison - 01.png |
| Figure 3.5‑5: stock trajectory |