**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**

**Tasks**

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 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
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**

In order to assess the impact of the changes made on the assessment (data used or model configuration), a series of diagnostics can be inspected :

1. model parameters :
2. model likelihood :
3. residuals for the observations :
4. model uncertainty
5. model stability : retro

**Approaches**

tasks 1) and 2) are simple, these new data have to be used and we just need to assess which difference they make on the assessment. The work just consists in running the model with updated data and describing the changes in model parameters and estimated quantities (model from HAWG2017 used as the base case).

For task 3) we have to decide whether the additional age-classes can be used in the assessment. I propose that we do that in 2 steps :

1. first fit the model using the additional age-groups for the IBTS with the default model parameters configuration (and compare with HAWG2017),
2. see if the model can be improved by changing the configuration (grouping) on observation variances and catchability of this survey

step a) could be based on comparing parameter values too see if model improves, comparing model uncertainty, retrospective pattern, residuals pattern... . step b) could be based on statistical tests (since we would compare model based on the same data).

For task 4) we can maybe use the same diagnostics as for task3/step a, but interpretation may be more difficult as the model will be quite different, with the introduction of time varying proportions per component.