

**This is an electronic reprint of the original article.**

**This reprint *may differ* from the original in pagination and typographic detail.**

**Author(s):** Anders Nielsen, Jukka Pönni, Luis Ridao Cruz, Massimiliano Cardinale, Noël Holmgren, Pekka Jounela, Zeynep Pekcan Hekim & Ruth Fernández

**Title:** Report of the Inter-benchmark Process on Herring (*Clupea harengus*) in the Gulf of Bothnia (IBPCLuB)

**Year:** 2019

**Version:** Published version

**Copyright:** International Council for the Exploration of the Sea 2019

**Rights:** CC BY-NC 4.0

**Rights url:** <http://creativecommons.org/licenses/by-nc/4.0/>

**Please cite the original version:**

ICES. 2019. Report of the Inter-benchmark Process on Herring (*Clupea harengus*) in the Gulf of Bothnia (IBPCLuB), 19–21 November 2018, by correspondence. ICES CM 2018/ACOM:67. 16 pp.

All material supplied via *Jukuri* is protected by copyright and other intellectual property rights. Duplication or sale, in electronic or print form, of any part of the repository collections is prohibited. Making electronic or print copies of the material is permitted only for your own personal use or for educational purposes. For other purposes, this article may be used in accordance with the publisher's terms. There may be differences between this version and the publisher's version. You are advised to cite the publisher's version.

# ICES IBPCLUB REPORT 2018

ICES ADVISORY COMMITTEE

ICES CM 2018/ACOM:67

## Report of the Inter-benchmark Process on Herring (*Clupea harengus*) in the Gulf of Bothnia (IBPCLuB)

19 – 21 November 2018

by correspondence



**ICES**  
**CIEM**

International Council for  
the Exploration of the Sea

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](http://www.ices.dk)  
[info@ices.dk](mailto:info@ices.dk)

Recommended format for purposes of citation:

ICES. 2019. Report of the Inter-benchmark Process on Herring (*Clupea harengus*) in the Gulf of Bothnia (IBPCLuB), 19–21 November 2018, by correspondence. ICES CM 2018/ACOM:67. 16 pp.

The material in this report may be reused using the recommended citation. ICES may only grant usage rights of information, data, images, graphs, etc. of which it has ownership. For other third-party material cited in this report, you must contact the original copyright holder for permission. For citation of datasets or use of data to be included in other databases, please refer to the latest ICES data policy on the ICES website. All extracts must be acknowledged. For other reproduction requests please contact the General Secretary.

The document is a report of an Expert Group under the auspices of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council.

## Contents

---

<b>Executive summary .....</b>	<b>1</b>
<b>1 Introduction .....</b>	<b>2</b>
1.1 Terms of Reference .....	2
1.2 Description of the Benchmark Process .....	2
<b>2 Gulf of Bothnia Herring (SD 3031).....</b>	<b>4</b>
2.1 Issue list.....	4
2.2 Estimate statistical conversion factors between day and night acoustic survey abundance indices (ToR a1) .....	4
2.3 Investigate selection pattern assumptions and other configuration parameters and final assessment (ToR a2) .....	6
2.4 Short term projections .....	9
2.5 Appropriate Reference Points (MSY).....	9
Proposed reference points.....	12
<b>3 Future Research and data requirements .....</b>	<b>14</b>
<b>4 External Reviewers Comments .....</b>	<b>14</b>
<b>5 Conclusions .....</b>	<b>15</b>
<b>6 New Stock Annex .....</b>	<b>15</b>
<b>7 References .....</b>	<b>15</b>
<b>Annex 1: List of participants.....</b>	<b>16</b>



## Executive summary

---

The inter-benchmark for Gulf of Bothnia Herring SD 30-31 was held by correspondence during 19–21 November 2018. The aim for the inter-benchmark was to evaluate the present analytical assessment method of herring with emphasis on the estimated statistical conversion factors between day and night of the acoustic survey abundance indices and to improve the assessment model settings by investigating selection pattern assumptions and other configuration parameters.

The working group tested the potential underestimation bias in the acoustic survey target strength (TS) caused by diel vertical migration patterns of herring between day and night times. This underestimation bias has been shown to underestimate abundance indices in the southern and western Baltic, where fish are close to the bottom during daytime and therefore not detectable with echosounder. The analyses suggested that diel vertical migration patterns are not a major issue in the abundance estimation of the Gulf of Bothnia stock and can be left out from the stock assessment considerations.

After the 2018 WGBFAS meeting and just before the start of ADGBS in May 2018 a mistake was discovered in one year of the survey input data for assessment of Herring in Gulf of Bothnia (GoB) in Sub-Divisions 30 and 31. The assessment run including the corrected data resulted in poor residual patterns and Mohn's rho values which were considered not acceptable. A pre-meeting was undertaken on 24<sup>th</sup> October 2018 during which preparatory work was agreed. On 15<sup>th</sup> November an updated assessment addressing the assessment model settings was presented. The configuration setup was revised and sensitivity runs were made by changing the configuration until finally finding the configuration setup that gave the lowest AIC values. The final assessment with improved configuration setup was approved during the video meeting on 20<sup>th</sup> November which can be reviewed under `gobherring_2018` in `stockassessment.org`. New reference points were calculated based on the new approved assessment and short term projections were given.

## 1 Introduction

---

After the 2018 WGBFAS meeting and just before the start of ADGBS in May 2018 a mistake was discovered in the input data for assessment of Herring in Gulf of Bothnia (GoB) in subdivisions 30 and 31. The year 2015 SD 30 acoustic index-values differed significantly from the ones issued by ICES WGBIFS and it was revealed that they had been wrong since the last Benchmark assessment in WKBALT (ICES, 2017b), where the mistake was traced down to.

A new run with corrected input data and forecast were made with the state space assessment model (SAM), which is used in the GoB herring stock assessment. However, the residuals and Mohn's rho values were not considered acceptable in this new run. This was due to the configuration that was initially set to fit the data which was not correct. A second new run with slightly adjusted configuration of SAM was also performed to compare the model outputs. (ICES 2018, WGBFAS report Annex 08: Survey input issue on Herring in Gulf of Bothnia).

During the ADGBS meeting the ACOM decided that an Inter-benchmark Process was needed to solve this issue. Since the process was already going to be held, it was decided to add to the benchmark process another issue, which came up during the 2018 WGBFAS meeting, i.e. estimation of statistical conversion factors in acoustic survey abundance indices between day- and night time.

### 1.1 Terms of Reference

Inter-benchmark process (IBP) on herring (*Clupea harengus*) in the Gulf of Bothnia (IBPCLUB), chaired by ICES Chair Noël Holmgren, Sweden, and attended by the invited external expert Luis Ridao Cruz, Faroe Islands, was established and met by correspondence on the 19–21 November 2018 to:

- a) Evaluate the present analytical assessment method of herring with emphasis on:
  1. Estimate statistical conversion factors between day and night acoustic survey abundance indices
  2. Improve assessment model settings:
    - i. Investigate selection pattern assumptions and other configuration parameters;
- b) Update the stock annex as appropriate;
- c) Re-examine and update MSY and PA reference points according to ICES guidelines (see Technical document on reference points);
- d) Prioritize recommendations for future improving of the assessment methodology and data collection.

### 1.2 Description of the Benchmark Process

The meeting was held by correspondence and scheduled for the 19–21 November. The acoustic data was made available from 6th October. On 22nd October it was clear that the acoustic data was not of the structure that TOR a1 could be resolved. A pre-meeting was undertaken on 24th October during which preparatory work was agreed. On 15th November an updated assessment addressing ToR a2 was presented. The actual meeting started as planned the 19th, but without the reviewer. The assessment was discussed on the 19th, and few alternative settings were proposed to be run until the next day. The group reconvened on the 20th, this time with the reviewer. The assessment was presented and accepted, after which the group could proceed with the calculation of the reference

points. A working document on the calculated reference points was uploaded to the SharePoint on the 23rd. A meeting to discuss the document was held on the 28th, during which minor comments were raised. The entire material was now ready to be written down in the report.



## 2 Gulf of Bothnia Herring (SD 3031)

### 2.1 Issue list

<u>Issue</u>	<u>Problem/Aim</u>	<u>Work needed / possible direction of solution</u>	<u>Data needed to be able to do this: are these available / where should these come from?</u>	<u>External expertise needed at benchmark</u> <u>type of expertise / proposed names</u>
(New) data to be Considered and/or quantified				
Tuning series				
Discards				
Biological Parameters				
Assessment method	The state space assessment model (SAM) (ICES WGMG report 2009) is used in the update assessment.	Adjust configuration of SAM model to produce acceptable residuals, retrospective patterns, Mohn's rho values and log-likelihoods.	No new data needed.	External expertise is needed in SAM configuration. Suggestion for expert: Anders Nielsen (DTU Aqua, DEN)
Biological Reference Points	Problem/Aim is to assess reference points for Her27.3031 stock after acceptable SAM configuration has been set	Use of flr and msy packages in R.	The data will be provided by SAM after acceptable configuration has been set.	External expertise is needed in ref.points assessment. Suggestion for expert: Massimiliano Cardinale (SLU, SWE)

### 2.2 Estimate statistical conversion factors between day and night acoustic survey abundance indices (ToR a1)

The working group tested the potential underestimation bias in acoustic survey target strength (TS) caused by diel vertical migration patterns between day and night times. This underestimation bias has been shown to underestimate abundance indices in the southern and western Baltic, where fish are close to the bottom during daytime and therefore not detectable with echosounder (ICES, 2017a; Orłowski, 2000, 2001, 2005). In the Gulf of Bothnia this potential underestimation bias of daytime target strength has not been taken into account in the abundance estimation even though the acoustic surveys of herring are conducted both during day and night time. Therefore, the aim of this assessment study was to estimate whether daytime TSs are different than that during night times and,

whether daytime TSs should be multiplied by an estimated multiplier to obtain unbiased estimates of abundance for daytime TS.

The acoustic survey data from 2007 – 2008 and 2011 – 2017 were used. The remaining acoustic data obtained from experts 2009 – 2010 was not used here. That is because in year 2009 the TS data was depth aggregated (i.e. the sum of TS over all depth zones in each coordinate at time t) and in year 2010 depth information was missing.

The acoustic TS patterns were recognized using gradient boosted machines (GBM, Friedman 2001). A GBM model was used here because the TS function estimation/approximation was viewed from the perspective of numerical optimization in function space, rather than parameter space. The parameters of GBM model were estimated using 10-fold cross-validation (Kohavi 1995) i.e. by partitioning test phase into 10 disjoint non-overlapping subsets and then, using all data once after finding the best parameters. The statistical analyses were done using RapidMiner software (version Studio Large 9.0.003, Mierswa *et al.* 2006).

The results suggested annual variation in diel vertical migration patterns. Depending on the year, the average predicted TS densities during the night times vs. daytimes were either higher or lower with no clear inter annual pattern (Figure 1).

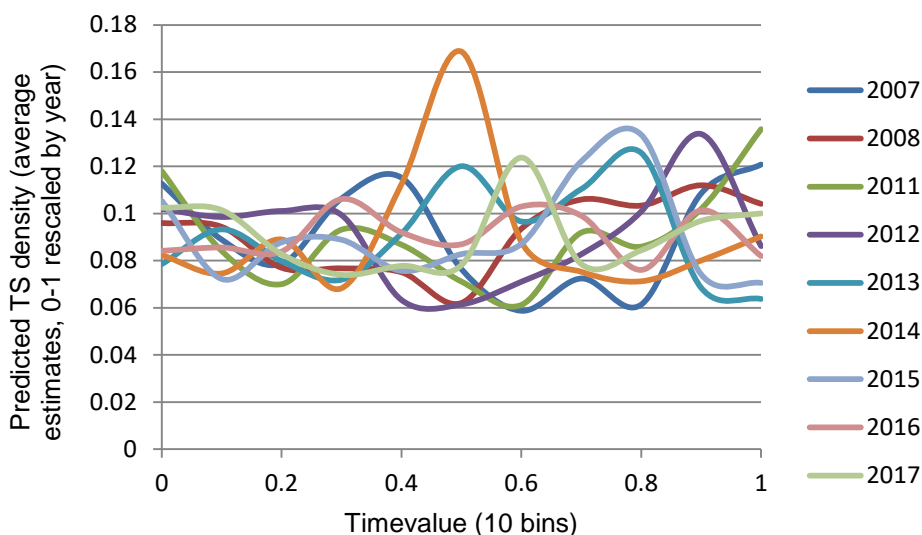


Figure 1. Predicted average annual TS density by timevalues 0 - 1 (times of the day, 0-24) in 2007-2008 and 2011 – 2017.

The average TS densities were random, which suggests that diel vertical migration patterns are not a major issue in the abundance estimation of the Gulf of Bothnia herring stock and can be left out from the stock assessment considerations.

A probable underestimation issue of Gulf of Bothnia herring abundance could relate to the predicted average depth dependent TS density that seems to vary a lot between the years. The predicted average depth dependent TS density was lower in upper water layers in years 2007 – 2008, 2013, 2016 and especially in 2017 than that in the other years (Figure 2). In these years the pelagic trawl may not have caught adequate numbers of fish even in the upper water depth zones. For example in 2017 the average towing depth of the pelagic trawl was 32 m.

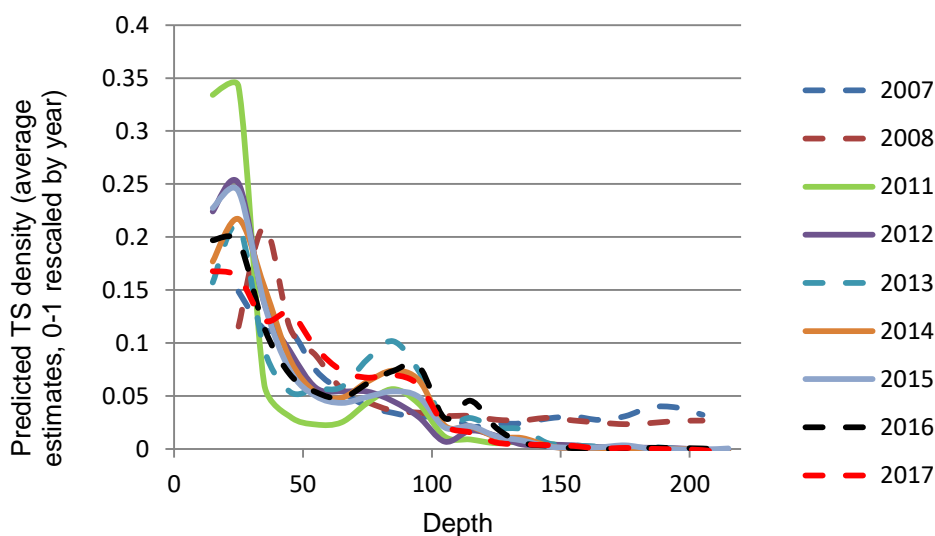


Figure 2. The average predicted TS density in water depths 15 – 215 m in years 2007 – 2008 and 2011 – 2017.

The IBPCLuB recommends the Baltic International Fish Survey Working Group (WGBIFS) to evaluate whether the annual variation in the predicted average TS density patterns in different water depths (Figure 2) impact the survey numbers that are used in the Gulf of Bothnia herring stock assessments.

### 2.3 Investigate selection pattern assumptions and other configuration parameters and final assessment (ToR a2)

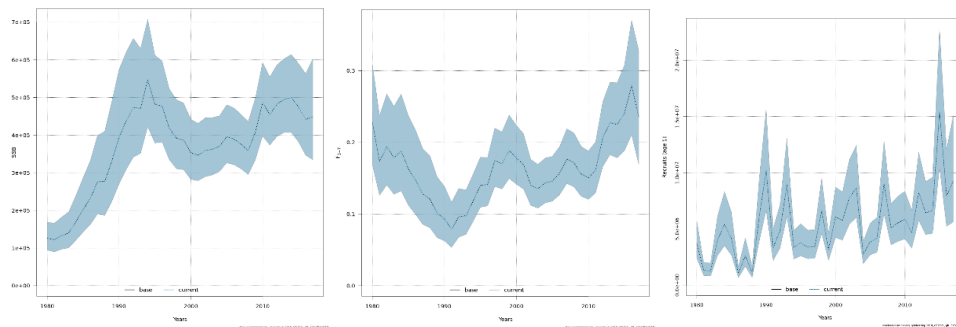
Following the Terms of Reference (a2), in order to find the configuration that would produce the best fit for the assessment model with the new data we carried out sensitivity runs. These runs were carried out in a step-wise manner, starting from the old configuration setup which had a poor fit in terms of AIC values (473.54) and also produced biased retrospective patterns. In each step we modified the section of the configuration and followed the outcomes in terms of AIC and logLikelihood values. During these step wise runs we kept the configuration setting that provided an improvement in the AIC values and applied the following change in the configuration. These stepwise changes can be found in Table 1. On the final run we only included the configuration changes (a, c, f see Table 1) that provided the best model performance in terms of 128 units lower AIC estimates compared to the model with the old configuration. For details please see WD2.

Table 1. Model results from the step-wise changes made starting from the old configuration to the new model configuration:

	Model/Change	log(Likelihood)	# parameters	AIC
	Benchmark configuration (old)	-221.7713	15	473.5426
a)	Correlated random walks for fishing mortality	-164.7481	16	361.4961
b)	Catchability more flexible	-158.4887	27	370.9773
c)	Single variance parameter for fishing mortality process	-158.8851	26	369.7703

d)	Single variance parameter for the survival process	-163.9815	25	377.9629
e)	Single observation variance parameter for each fleet	-174.5269	21	391.0538
f)	AR(1) correlation structure for survey observations (this is the same as the “new” WGBFAS 2018 suggested configuration)	-162.8125	23	371.6250
a), c), and f)	Inder-benchmark suggested (see appendix B)	-155.8154	17	345.6308

The final configuration that gave the best fit was included in the assessment model and the assessment can be viewed under the run Gobherring\_2018 in stockassessment.org. The final assessment plots for SSB, F and Recruitment can be found in Figure 3. The final year estimates for SSB, F and Recruitment differed by 4%, 6% and 31% compared to the final assessment estimates from the assessment run “RevisedHer30312018” which was the assessment 2018 with the old configuration. The residuals from the run with the new configuration (Gobherring\_2018) also improved compared to the old run especially the 2015 acoustics is improved (Figure 4). The Mohns rho values in the final assessment model for SSB, F and recruitment are 0.22, -0.17 and 0.50 respectively (Figure 5). This was an improvement from Mohns rho values for SSB 0.24, F 0.19 and Recruitment 0.71 from the assessment run made with the old configuration (including the correct data).



**Figure 3. Output of SSB, F and recruitment from the Gobherring\_2018 including the new improved configuration.**



Figure 4. Residuals from the Gobherring\_2018 including the new improved configuration.

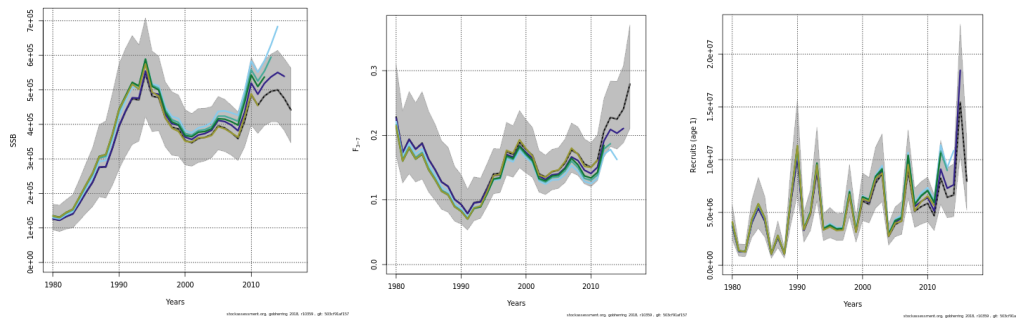


Figure 5. Retrospectives from the Gobherring\_2018 including the new improved configuration. The Mohns rho 0.22, -0.17 and 0.50

During the analyses it was realised that the estimated number at age 1 in some years were smaller than the estimated number of fish at age 2 in the following year. This is probably due to the age 1 fish inhabiting somewhat different areas than age 2 fish. The acoustic survey is done offshore and is probably not able to detect the age 1 fish that are inhabiting more inshore areas while the age 2 fish are better represented in the areas that the acoustic survey covers. It could also be due to the mixing of the two stocks (Gulf of Bothnia and Gulf of Finland stocks) that spawn in the same area in the Archipelago Sea.

During the inter-benchmark WebEx meeting it was also suggested to evaluate the impact of density dependence in the trap-net survey. The density dependence decreased model performance in terms of increased AIC estimates and thus, density dependence was not included into the final model.

## 2.4 Short term projections

The short term projections were run based on the new stock assessment (Table 2) and can be found in gobherring\_2018 in stockassessment.org.

**Table 2. Short term forecast based on the new stock decided at IBPCLuB.**

	CATCH (2019)	F <sub>TOTAL</sub> (2019)	SSB (2019)	SSB (2020)
ICES advise basis*				
Fmsy precautionary	107215	0.229	483943	453672
Fpa	109302	0.234	483578	450945
Flim	139253	0.309	477892	418241
Blim (2020)	336081	1.016	426305	199308
Bpa (2020)	261673	0.687	449440	279111
Btrigger (2020)	261673	0.687	449440	279111
Fmsy Upper	107215	0.229	483943	453672
Fmsy Lower	79012	0.164	488999	486234

- With 84 599 TAC in 2018

## 2.5 Appropriate Reference Points (MSY)

The reference points were also updated during the inter-benchmark.

**Table 3. Summary table of stock reference points before the inter-benchmark**

REFERENCE POINT	VALUE	TECHNICAL BASIS
Current F <sub>MSY</sub>	0.21	Eqsim
Current B <sub>lim</sub>	202272	Eqsim
Current B <sub>pa</sub>	283180	Eqsim
Current MSY B <sub>trigger</sub>	283180	Eqsim

The analysis in this report uses the newest (1980-2017) assessment results from the IBPCLuB inter-benchmark SAM assessment (model: gobherring\_2018).

Eqsim was used for this stock. Settings for the Eqsim can be found in Table 4.

**Table 4. Settings used for the Eqsim**

DATA AND PARAMETERS	SETTING
SSB-recruitment data	Full data series
Exclusion of extreme values (option extreme.trim)	Not used

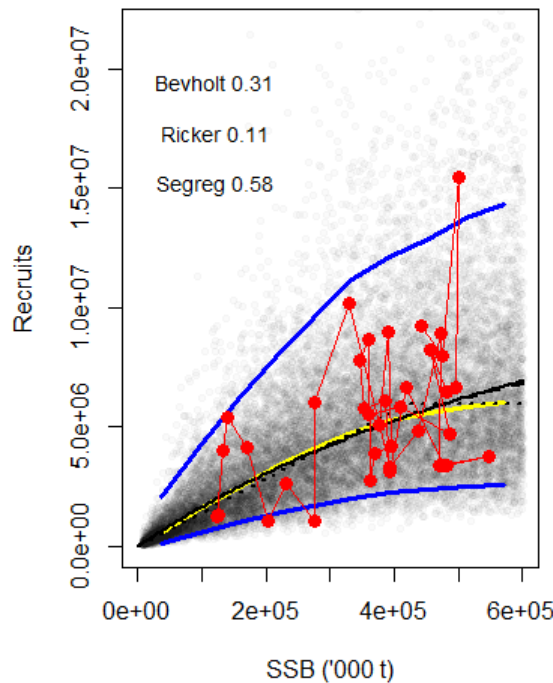
Mean weights, proportion mature and F at age pattern	2008–2017
Exploitation pattern	2008–2017
Assessment error in the advisory year. CV of F	0.212
Autocorrelation in assessment error in the advisory year	0.423

The stock recruitment fit using the three models (Ricker, B&H and segmented regression) weighted by the default "Buckland" method available in EqSim gave a "straight" line for all models (Table 5, Figure 6).

**Table 5. The parameter estimates and contribution of each of the initial models, which gave a "straight" line for all models shown in Figure 1.**

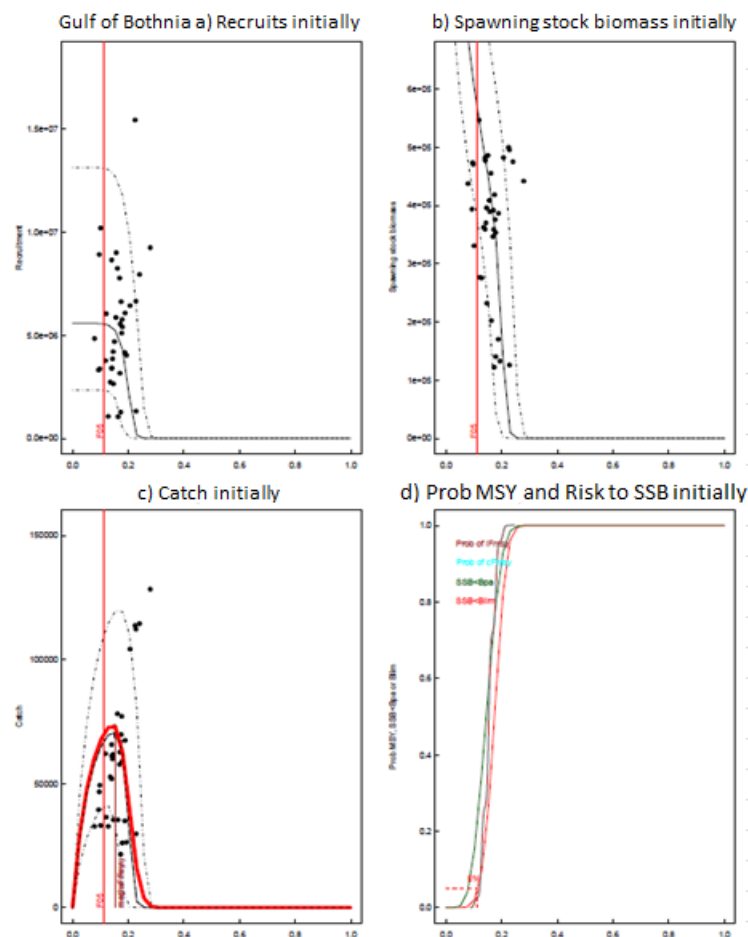
Model	a	b	cv	prop
Bevholt	18.29608	9.746547e-07	0.5270003	0.313
Ricker	17.81932	7.512193e-07	0.5269513	0.107
Segreg	14.15055	4.212513e+05	0.5272382	0.580

**Initial predictive distribution of recruitment for Gulf of Bothnia**



**Figure 6. The stock recruitment fit using the three models (Ricker, B&H and segmented regression) weighted by the default "Buckland" method available in EqSim gave a "straight" line for all models. The yellow and blue lines represent the median and 5% and 95% percentiles of the distributions of the stochastic recruits drawn from the models.**

Thus, a segmented regression model was used with a breakpoint set arbitrarily at the average observed SSB (i.e.  $B_{lim} = 368\,244\text{ t}$ ) as dictated by ICES guidelines for reference point estimation (ICES, 2017c). However, this resulted in an unrealistically large value of  $B_{pa}$  (471 300 t) and thus in an unrealistically low value of  $FP_{0.05}$  (5% risk to  $B_{lim}$ ; 0.112; Figure 7).



**Figure 7. The initial EqSim model simulation suggested 95% risk of overexploitation in 33 years (out of 38 years in total) even though SSB approx. four-folded from 1980 to 2017. This simulation was considered as implausible and hence ICES reference points guidelines were modified.**

Thus, the ICES reference points guidelines were modified as follows; the first step was to estimate  $F_{MSY}$  using a hockey stick SR relationship with  $B_{lim}$  at the average SSB and without  $MSY\ B_{trigger}$ , but with assessment and advice error (i.e. using the default values). Once the  $F_{MSY}$  was estimated, the simulations were run again with the same hockey stick SR relationship and  $B_{lim}$  to estimate  $MSY\ B_{trigger}$  defined as the 5th percentile of the SSB at  $F_{MSY}$ . Successively,  $B_{pa}$  was set as  $MSY\ B_{trigger}$  and a new value of  $B_{lim}$  was estimated as  $B_{pa}$  divided by  $\exp(1.645 \times 0.2)$ . After  $B_{lim}$ ,  $B_{pa}$  and  $MSY\ B_{trigger}$  were all defined, the ICES procedure for setting the reference points was used to estimate the remaining reference points. The SR relationship used for these runs was a hockey stick with the breakpoint set at the new  $B_{lim}$ . The number of samples used to fit the SR relationship and the number of runs used in all EqSim simulations were 1000 and 200, respectively. Autocorrelation of recruitment was used in all EqSim simulations.  $F_{pa}$  was estimated using the ICES standard procedure ( $F_{pa} = F_{lim} \times \exp(-1.645 \times \sigma)$ ). Sigma was estimated as the uncertainty associated to the F in last year of the assessment (i.e. 2017;  $\sigma = 0.150$ ).

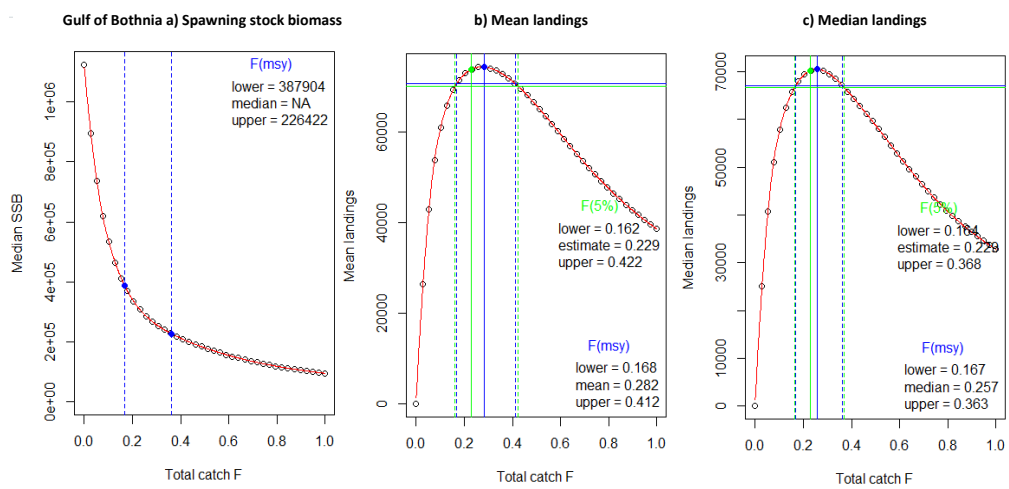


**Proposed reference points**

Summary table of proposed stock reference points:

REFERENCE POINT	VALUE
FP.05 (5% risk to B <sub>lim</sub> ) with MSY B <sub>trigger</sub>	0.23
FP.05 (5% risk to B <sub>lim</sub> ) without MSY B <sub>trigger</sub>	0.21
F <sub>MSY</sub>	0.26
F <sub>MSY</sub> precautionary	0.23
F <sub>MSY</sub> lower	0.167
F <sub>MSY</sub> upper	0.36
F <sub>pa</sub>	0.23
F <sub>lim</sub>	0.31
F <sub>MSY</sub> upper precautionary	0.23
F <sub>MSY</sub> range with MSY B <sub>trigger</sub>	0.164–0.23
F <sub>MSY</sub> range without MSY B <sub>trigger</sub>	0.156–0.21
MSY B <sub>trigger</sub>	279 110 t
B <sub>pa</sub>	279 110 t
B <sub>lim</sub>	199 364 t

As explained above, the standard ICES procedure for setting the Blim reference point in this case would result in an unrealistically large value of Blim and thus in an unrealistically low value of FP0.5. The SR relationship does not show any density dependence and hence it is difficult to justify the exact FMSY level. Thus, the procedure used to estimate the reference points for herring in SD 30 and 31 is not in strictly in accordance with the ICES reference points guidelines but it has been modified to account for the specific SR relationship of this stock. Also, according to the EqSim estimations, FP0.5 (0.229) is lower than FMSY (0.257) estimated with MSY Btrigger (Figure 8) and thus FMSY and the FMSY range are dictated by precautionary considerations in this case; FMSY and FMSY upper are capped by FP0.5 to 0.229 (and rounded to 0.23).



**Figure 8. Summary plots of FMSY range for Herring in Subdivision 30 and 31 with MSY Btrigger.**

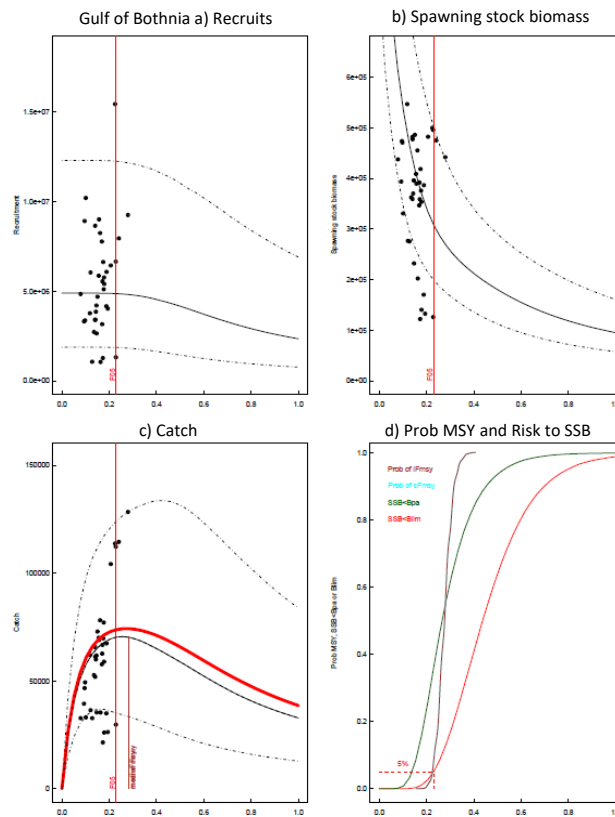


Figure 9. EqSim results for Herring in Subdivision 30 and 31 with MSY Btrigger.

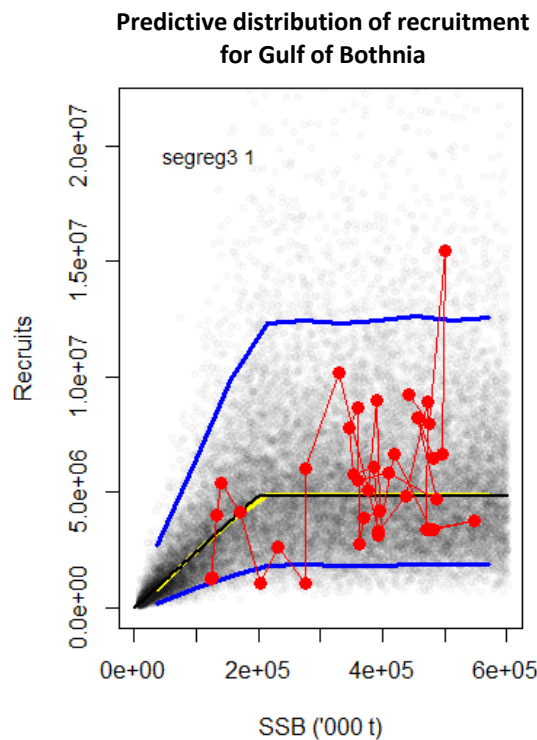


Figure 10. Stock recruitment relationship (i.e. segmented regression with breakpoint at Blim) for Herring in Subdivision 30 and 31 used in the EqSim simulations for the estimation of the MSY reference points. The yellow and blue lines represent the median and 5% and 95% percentiles of the distributions of the stochastic recruits drawn from the final model.

### 3 Future Research and data requirements

---

In the last Benchmark (WKBALT) in 2017, it was recommended

- 1) to consider genetic studies between the areas, and tagging studies to provide supporting information for the combination or separation since there is no strong biological evidence either for combining or separating SDs 30 and 31 for stock assessment.
- 2) to consider the possibilities to the extension of the acoustic survey to the suitable parts (i.e. deep enough waters in southern/middle parts) of SD 31.

These recommendations are still valid.

As mentioned in section 2.2. the IBPCLuB recommends the Baltic International Fish Survey Working Group (WGBIFS) to evaluate whether the annual variation in the predicted average TS density patterns in different water depths (Figure 2) impact the survey numbers that are used in the Gulf of Bothnia herring stock assessments.

As mentioned in section 4. there are concerns about the relatively large retrospective pattern in both SSB and F which suggests that the assessment model overestimates the herring stock. These are issues that need further investigation in future benchmarks.

### 4 External Reviewers Comments

---

The stock was re-evaluated with the same assessment model (SAM) but with modified configuration options. The assessment and evaluation of reference points followed the stock annex for Gulf of Bothnia Herring SD 30–31.

The resulting assessment improved the overall fit to the data with lower standardized one-observation-ahead residuals and fewer blocks of both positive and negative residuals. Retrospective analysis suggest overestimation of SSB and consequently and underestimation of average fishing mortality (F3-7). Just one of the retrospective runs fall out of the uncertainty bands of the adopted assessment.

The stock increased substantially from 1980 to mid-1990's due to lower catches in the 1980's. From 1990 to 2000 catches raised two-fold from 30 000 t. to 60 000 t. causing the stock to drop considerably to 350 000 t. Although catches have increased to historical levels since 2010, SSB has remained stable at around 470 000 t. as a consequence of higher than average recruitment (5.4 mill.). Estimated SSB was only below  $MSY B_{trigger}$  from 1980 to 1988. Estimated fishing mortality has been above  $F_{MSY}=0.229$  since 2015.

Biological reference points were evaluated with the updated assessment output. The procedure followed the previous benchmark directives and it resulted in upwards revision of  $F_{MSY}$  from 0.21 to 0.23 in IBPCLuB\_2018.  $MSY B_{trigger}$  decreased from 283 180 t. to 279 110 t.

The reviewer confirms that the outcomes of the benchmark are appropriate to provide scientific advice.

Since 2010 the stock has remained stable at around 473 000 t. even though fishing mortality was higher than  $F_{MSY}=0.23$  from 2015 to 2017. There are concerns about the relatively large retrospective pattern in both SSB and F which suggests that the assessment model overestimates the herring stock. These are issues that need further investigation in future benchmarks.

## 5 Conclusions

---

The IBPCluB working group and the reviewer agree that the outcomes of this benchmark process are appropriate to provide scientific advice.

## 6 New Stock Annex

---

The new Stock Annex can be found here:

[http://community.ices.dk/ExpertGroups/StockAnnexes/Stock\\_Annexes/her.27.3031\\_SA.docx](http://community.ices.dk/ExpertGroups/StockAnnexes/Stock_Annexes/her.27.3031_SA.docx)

## 7 References

---

- Friedman, J.H. 2001. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics* 2001, Vol. 29, No. 5, 1189–1232
- ICES. 2017a. Manual for the International Baltic Acoustic Surveys (IBAS). Series of ICES Survey Protocols SISP 8 - IBAS. 47 pp. <http://doi.org/10.17895/ices.pub.3368>
- ICES. 2017b. Report of the Benchmark Workshop on Baltic Stocks (WKBALT), 7–10 February 2017, Copenhagen, Denmark. ICES CM 2017/ACOM:30. 108 pp.
- ICES. 2017c. ICES fisheries management reference points for category 1 and 2 stocks. ICES Advice Technical Guidelines. DOI: 10.17895/ices.pub.3036
- ICES. 2018. Report of the Baltic Fisheries Assessment Working Group (WGBFAS), 6–13 April 2018, ICES Headquarters, Copenhagen, Denmark. ICES CM 2018/ACOM:11.
- Kohavi R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *IJCAI'95 Proceedings of the 14th international joint conference on artificial intelligence*. Morgan Kaufmann Publishers, San Francisco, CA, p 1137–1143
- Mierswa, I., Wurst, M., Klinkenberg, R., Scholz, M., Euler, T. 2006. Yale: rapid prototyping for complex data mining tasks. In: *Proceedings of the 12th ACM SIGKDD international conference on knowledge discovery and data mining (KDD '06)*. ACM, New York, NY, p 935–940
- Orłowski, A. 2000. Diel dynamics of acoustic measurements of Baltic fish. *ICES J. Mar. Sci.* 57: 1196–1203.
- Orłowski, A. 2001. Behavioural and physical effect on acoustic measurements of Baltic fish within a diel cycle. *ICES J. Mar. Sci.* 58: 1174–1183.
- Orłowski, A. 2005. Experimental verification of the acoustic characteristics of the clupeids daily cycle in the Baltic. *ICES J. Mar. Sci.* 62: 1180e1190.

## Annex 1: List of participants

---

Name	Institute	Country	Email
Anders Nielsen	DTU Aqua	Denmark	an@aqua.dtu.dk
Jukka Pönni	Natural Resources Institute	Finland	Jukka.Ponni@luke.fi
Luis Ridao Cruz (reviewer)	Marine Research Institute	Faroe Islands	luisr@hav.fo
Massimiliano Cardinale	SLU Aqua	Sweden	massimiliano.cardinale@slu.se
Noél Holmgren (Chair)	SLU Aqua	Sweden	noel.holmgren@slu.se
Pekka Jounela	Natural Resources Institute	Finland	pekka.jounela@luke.fi
Zeynep Pekcan Hekim	SLU Aqua	Sweden	zeynep.pekcan.hekim@slu.se
Ruth Fernández	ICES secretariat	other	ruth.fernandez@ices.dk

---