INTER-BENCHMARK PROTOCOL ON CELTIC SEAS WHITING (IBPCSWHITING 2021)

VOLUME 3 | ISSUE 103

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 3 | Issue 103

INTER-BENCHMARK PROTOCOL ON CELTIC SEAS WHITING (IBPCSWHITING)

Recommended format for purpose of citation:

ICES. 2021. Inter-Benchmark Protocol on Celtic Seas Whiting 2021 (IBPCSWhiting).
ICES Scientific Reports. 3:103. 31 pp. https://doi.org/10.17895/ices.pub.8718

Editors

Sofie Nimmegeers

Authors

David Stokes • Hans Gerritsen • Sara-Jane Moore • Sofie Nimmegeers

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

Contents

i Executive summary

IBPCSWhiting was primarily tasked with reviewing the data raising procedures after an error was identified during the Celtic Seas working group (WGCSE) in May, causing a subsequent shift in SSB in the early time-series. As this had potential implications for the reference points, the group also updated the stock's reference points. In preparation of the IBPCSWhiting, a review of ICES Accessions data revealed that a small amount of discard data for Belgium for data year 2019 (79 t) and 2020 (65 t) was available as well as 4.6 t for Spain in data year 2019. These were imported into InterCatch and the IC files for data year 2019 and 2020 only were exported again for use in this IBP. Furthermore, an anomaly was identified in the stock mean weights at age file that had been included at the 2020 benchmark (WKCELTIC). Additional information for catch data prior to 2003 was not available to the WKCELTIC so the 1999–2002 part of the time-series could not be revised. The catch mean weights-at-age file that was used to generate the Jan 1st stock mean weights at age, contained stock weights in the 1999–2002 rows, while the remaining 2003 onward were correct. The catch and stock mean weights at age files have been corrected for use in this IBP. The assessment model, a State-space Assessment Model (SAM) that is available on stockassessment.org[1] was re-run with the revised catch data inputs (additional discard data, re-estimated weights at age, and a revision to the allocation of sampling across catch).

[1] https://www.stockassessment.org/setStock.php?stock=whg.7b-ce-k_IBP_final.

ii Expert group information

Expert group name	Inter-Benchmark Protocol on Celtic Seas Whiting (IBPCSWhiting 2021)
Expert group cycle	Annual
Year cycle started	2021
Reporting year in cycle	1/1
Chair	Sofie Nimmegeers, Belgium
Meeting venue and dates	23–24 September 2021, Online meeting (7 participants)

1 Introduction

Whiting (*Merlangius merlangus*) in divisions 7.b-c and 7.e-k (southern Celtic Seas and eastern English Channel) – whg.27.7b-ce-k

The Inter-benchmark Protocol on Whiting (*Merlangius merlangius*) in divisions 7.b-c and 7.e-k (southern Celtic Seas and western English Channel; IBPCSWhiting 2021) met by correspondence during two MS Teams meetings, chaired by Sofie Nimmegeers (Research Institute for agriculture, fisheries and food (ILVO), Belgium) and attended by invited external experts Joanne Morgan (DFO, Canada) and Jesper Boje (National Institute of Aquatic Resources, The Technical University-ty of Denmark (DTU Aqua), Denmark).

Celtic Sea whiting 27.7b-ce-k was benchmarked alongside Celtic Sea haddock and cod as part of WKCELTIC (ICES, 2020). A shared term of reference across the three stocks was, as far as practical, to standardize and document methodologies for data checking and raising of national catch data submitted to InterCatch. To this end an existing R markdown document was modified and shared across the three stock coordinators. This documents the checking of the fishery-dependent catch data as well as the raising process being used for, and since, WGCSE2020.

During the data preparation of Celtic Sea whiting for the 2021 meeting of WGCSE an error was identified in the whiting R code that raises un-sampled catch data to international catch numbers-at-age. This was rectified just prior to the 2021 meeting in May and presented in plenary for discussion. The revised time-series for WGCSE2021 (data years 1999–2020) was presented vs. the previous year's dataseries (1999–2019) used at WGCSE2020. Specifically, the issue resulted in some un-sampled catches not being correctly linked to a corresponding sample and therefore an underestimate in total catch numbers-at-age for that métier. This could be seen as a SOP difference between the Catch in Tonnes file from InterCatch and the sum of the raised catch numbers-at-age x catch mean weights at age file also extracted from InterCatch. Historically Celtic Sea whiting assessment had seen reasonably high SOP values so the differences due to this missed allocation were not immediately picked up on.

At WGCSE 2021 it was initially agreed to move forward with the assessment and include the corrected data given the updates related mostly to the early part of the time-series. However, the overall perception of early stock development in terms of SSB had changed significantly and thus likely to effect the reference points and advice. In that context it was therefore agreed at WGCSE 2021 in May that before being used as the basis for advice time should be given to re-checking the code in detail and revising reference points as required.

2 Approach

The main issue for this interbenchmark was to:

1. Re-run the InterCatch raising procedure including the non-sampled component of catch and tune the model to the data.
2. Re-examine (and update if necessary) MSY and PA reference points according to ICES guidelines (see Technical document on reference points).

Given a lot of changes to data compilation and raising had been implemented at the recent benchmark, as well as to the assessment approach and input files, it was obviously important to use the time to do a complete audit of the data files along with the R code being used to compile the final assessment inputs. This was carried out during August–September 2021. Two additional data anomalies were identified and corrected. Having identified any issues, the reference points were re-calculated with the corrected data using the same R code, settings and time-series (data years 1999–2019) as agreed and used in the WKCELTIC 2020 Benchmark.

3 Input data

A number of issues were addressed and are outlined below chronologically according to the sequence in which they were ameliorated.

3.1 Missing data

A review of ICES Accessions revealed that a small amount of discard data for Belgium for data year 2019 (79 t) and 2020 (65 t) was available as well as 4.6 t for Spain in data year 2019. These were imported into InterCatch and the IC files for data year 2019 and 2020 only were exported again for use in this IBP. Comparing files used at WGCSE 2020 with the new ones from IC showed no other changes beyond the additions above.

3.2 Stock weights

While auditing the full set of data files going into the assessment an anomaly was identified in the catch mean weights at age file that had been used at the 2020 benchmark. Additional information for catch data prior to 2003 was not available to the benchmark so the 1999–2002 part of the time-series could not be revised. Historic time-series data are therefore copied in and still used for these years including catch mean weights at age and stock weights at age. It was discovered that the 1999–2002 rows in the catch mean weights at age file were stock weights, while the remaining 2003 onward were correct. The catch mean weights at age file is used to generate the Jan 1st stock mean weights at age so obviously would have an effect for that early part of the time-series (see WD01[2]).

The catch and stock mean weights at age files have been corrected for use in the IBP final run on Stockassessment.org[3].

3.3 InterCatch raising procedure

Since the 2020 benchmark, the raising of national catch data uploaded to InterCatch to produce international catch numbers-at-age and catch mean weights at age is documented through an R markdown document. The process does routine exploratory checks, allocates discards where only landings are available and finally allocates numbers-at-age sampling where un-sampled catches are uploaded (see WD02[4]). A walk through the full process was done during this IBP with particular focus on the specific issue of how samples are allocated to un-sampled catches.

First step: Estimate discard weights (tonnage)

Apply a mean discard ratio to landings strata that have no discards associated with them (zero discards are assumed to be valid observations and will NOT be filled in)

There are three steps for filling in:

[2] https://community.ices.dk/ExpertGroups/benchmarks/2021/IBPCSWhiting.

[3] https://www.stockassessment.org/setStock.php?stock=whg.7b-ce-k_IBP_final.

[4] https://community.ices.dk/ExpertGroups/benchmarks/2021/IBPCSWhiting/2021_Meeting_Documents/04.%20Working%20documents/WD02_aggregate_IC_data_whg.27.7b-ce-k_Sept_2021.html.

a) Group by country, gear and year (ignore differences between quarters, areas, broad gear groups)
b) Group by gear and year (if you don't have any data above, also combine countries)
c) Year (for all remaining cases, combine all data for the year)

```
# First calculate mean discard ratio by country, gear and year
a <- subset(ld,DisData) %>% group_by(Country,Fleet1,Year) %>% summarise(DisRatio1=sum(Dis)/sum(Lan))
# Next calculate mean discard ratio by gear and year
a <- subset(ld,DisData & Fleet1!='MIS_MIS') %>% group_by(Fleet1,Year) %>% summarise(DisRatio2=sum(Dis)/sum(Lan))
# for fleets that have no discard data at all, use overall ratio per year
a <- subset(ld,DisData) %>% group_by(Year) %>% summarise(DisRatio3=sum(Dis)/sum(Lan))
```

Second step: fill in sample data (numbers-at-age)

First check which catch (landings or discard) records have missing sample data (so catch weight but no age data)

Join caton (weight) on canum (numbers-at-age) and identify where there is no match

There are only two steps for filling in sampling data:

a) Where there is no match, assume that each country's sampling is representative of the missing data of that country (in the relevant year, quarter)
b) Where a country has no sampling data, assume that the overall sampling data (all countries combined) is representative for that country (in the relevant year, quarter)

```
## Assume that each country's sampling is representative of the missing data of that country - VERY BROAD ASSUMPTION!
# mean lf by country, season and year
a <- canum1a %>% group_by(Year,Season,Country,Catch.Cat.,AgeLength) %>%  summarise(n=sum(Frequency),w=sum(Frequency*Weight)/sum(Frequency))
a <- canum1a %>% group_by(Year,Season,Catch.Cat.,AgeLength) %>% summarise(n=sum(Frequency),w=sum(Frequency*Weight)/sum(Frequency))
```

In allocating samples to un-sampled catches the initial task is to create a reference table of which fleets have sampling or not. We then determine the appropriate level of sample allocation based on step 2 a) or b; explained above) and enter a label into the sample allocation table accordingly.

The issue identified pre-WGCSE2021 was that this stage in the process was based on a general fleet aggregation (OTB-DEF, TBB-DEF, etc ('Fleet1 ')) rather than the more detailed métier level (OTB_DEF_70_99_0_0_All, TBB_DEF_70_99_0_0_All , etc ('Fleets ')). In effect where a sample appeared anywhere in the higher grouping of fleet, the code flagged that a sample was available and it was therefore not passed to the next level for sample allocation (Step 2b explained above). The subsequent actual allocation process however used the more granular relationship between catches and samples of the detailed métier name field ('Fleets'). The result being that a number of catches that shared a gear group ('Fleet1 ') and that had sampling flagged as available for that group, did not actually have a sample specific to that métier and the catch numbers-at-age of that métier were therefore missed (Figure 3.1).

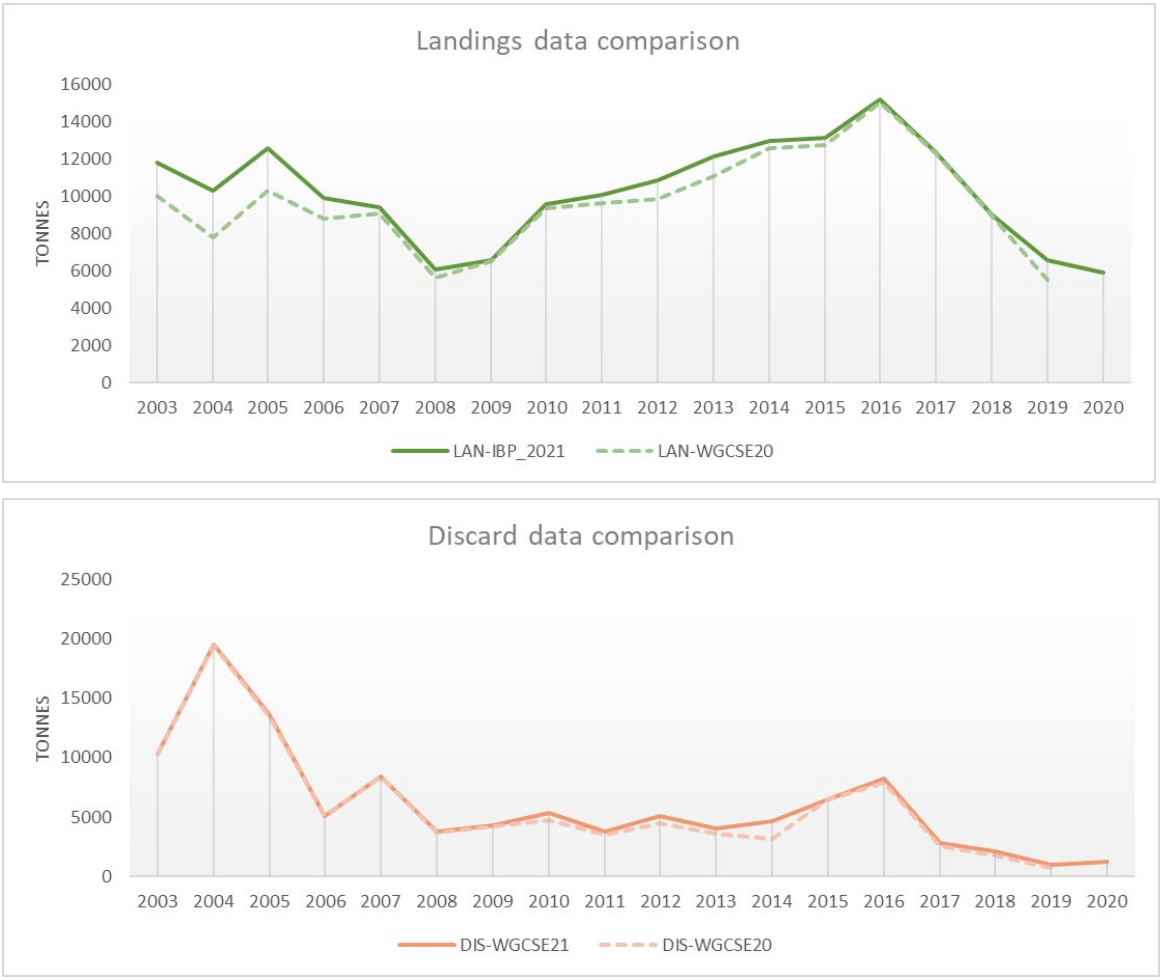


Figure 3.1. Comparison of raised international landings and discard data between assessment year 2020 and corrected data for assessment year 2021 used in the IBPCSWhiting .

The conclusion following a thorough discussion was that the table relationship had been identified and corrected based on appropriate SOP values. A further IBPCSWhiting discussion centred on whether it was appropriate to use only two steps in allocating the sample data; it would seem more logical to group by gear, country, year and season first before grouping by country, year and season and finally by year and season.

The implications for the estimated numbers-at-age were investigated and including gear in the first step did not change the estimates by more than 1% for the majority of years/ages (Figure 3.2). These differences due to potential catchability differences between fleets are considerably smaller than the inherent year-to-year differences in catch numbers-at-age (which regularly vary by more than 100% from one year to the next). Therefore it is unlikely that the choice of allocation method had an important effect on the catch numbers for this stock.

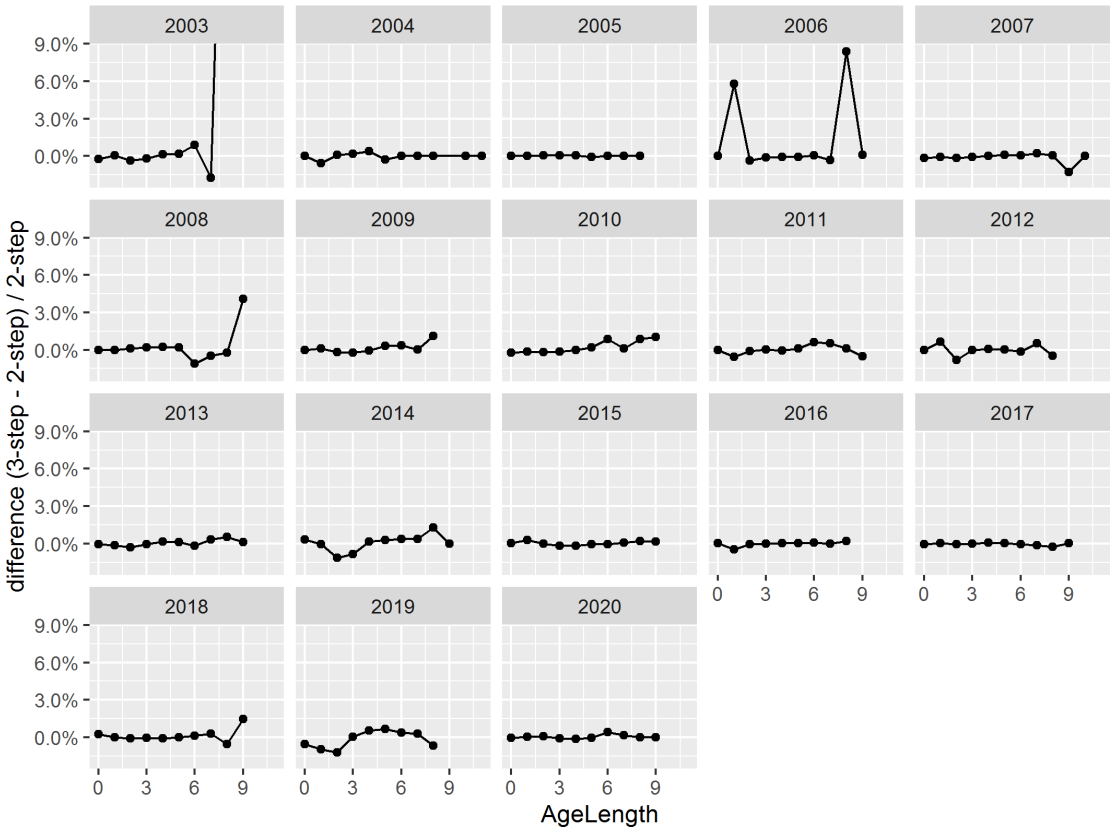


Figure 3.2 Difference between allocating un-sampled age compositions first by gear, county and season (3-step) or first by country and season (2-step).

4 Assessment

The existing SAM model as implemented in stockassessment.org, was re-run with the revised data inputs. For detailed configurations, please see Annex 3.The revision in SSB is notable from the early part of the time-series (Figure 4.1) and obviously a culmination in the stock weights error and unsampled catches issue for the years 2003–2006 in particular. Less noticeable differences in fishing mortality and recruitment were recorded.

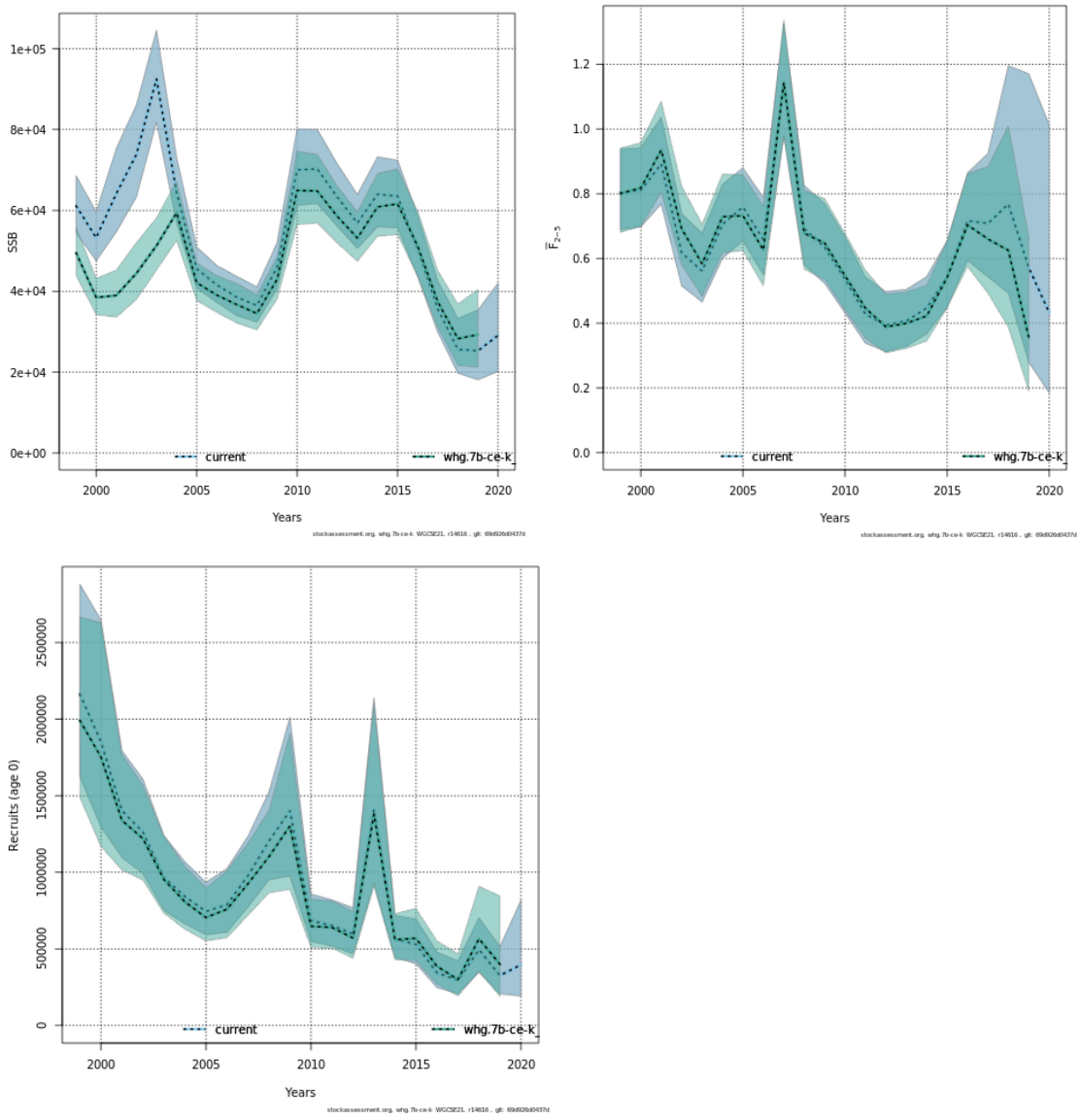


Figure 4.1. Summary plots from IBPCSWhiting (current – blue shading) compared with WGCSE 2020 (green shaded confidence intervals).

Comparing SSB estimates from the final IBPCSWhiting SAM run against the historic pre-benchmark XSA assessment at WGCSE 2019 shows much similar values following the corrections undertaken at this point (Figure 4.2).

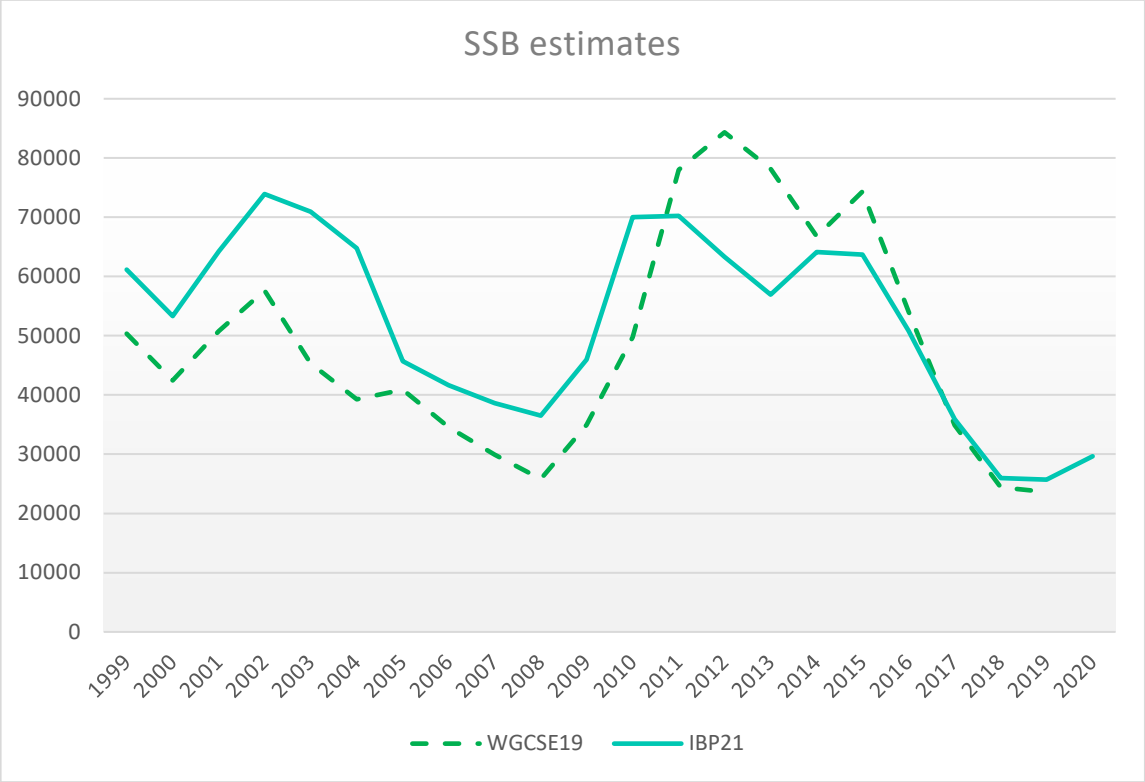


Figure 4.2. SSB estimates from the WGCSE2019 XSA assessment compared to IBPCSWhiting 2021.

Recruitment and fishing mortality are relatively unchanged and the group agreed the assessment reflected the likely changes expected from the updates to the input data. The final assessment run that was accepted by the group is in stockassessment.org[5].

It was agreed this assessment run would be cloned and used to generate the revised MSY and PA reference points.

Retrospective analyses of the updated assessment model shows an increase in retrospective patterns compared to the previous year's assessment, but the consistency of estimated SSB remains within acceptable limits (Figure 4.3). However, a number of options were discussed and tested in terms of the retrospective bias, including a run with the commercial LPUE index removed, but the decision was to leave the LPUE in at this time .

Before the IBP the developers of Stockassessment.org also explored settings to improve the bias, but with minimal improvement. All configuration settings therefore remain as per WKCELTIC 2020 and WGCSE 2019.

[5] https://www.stockassessment.org/setStock.php?stock=whg.7b-ce-k_IBP_final.

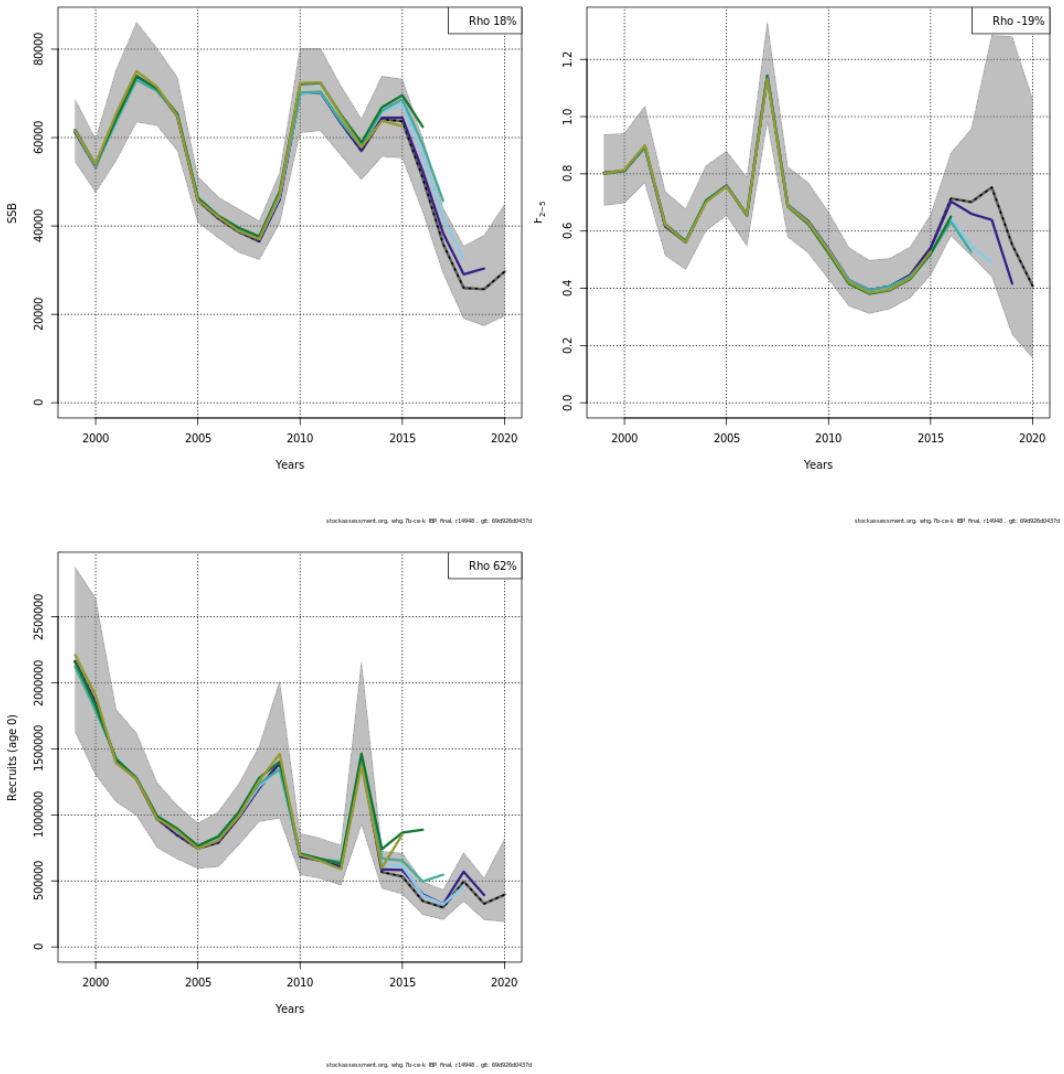


Figure 4.3. Retrospective runs at IBPCSWhiting 2021 for Celtic Sea whiting giving Mohn's Rho values for bias.

5 Reference points

In order to evaluate changes to the reference points due to revised data only, the final IBP SAM assessment run was cloned in Stockassessment.org so the 2020 data year could be removed without affecting the final assessment. That run is otherwise identical with the final IBP assessment run and available in stockassessment.org with the name: whg.7b-ce-k_IBP_RefPts[6]. It produces the stock object used as input to the reference points calculation, detailed in the R markdown working document 3[7]. The estimation of reference points follows the ICES Reference Points Guidance (ICES, 2021), using the R-programme EqSim.

Discussion was held at IBPCSWhiting around selection of the most appropriate stock type. WKCELTIC 2020 concluded with type 5 for whiting where there is no obvious impaired recruitment or link between S-R. It was discussed whether type 3 would be appropriate? The point was made that in general Celtic Sea gadoid stocks have been seen to recover from very low levels and therefore impaired recruitment can be more subjective with these stock. Also, fishing mortality is known to have been high historically and not included in the current time-series so it would be more appropriate to include reliable historic data if applying the type 3 approach to ensure SSB values extend over that period. In the absence of a definitive argument to the contrary, the group decided to maintain the WKCELTIC 2020 approach.

The S-R model fit changed somewhat compared to WKCELTIC 2020 with the segmented regression fit increasing from 0.598 to 0.787 (Figure 5.1.). The Segreg breakpoint now being 53 454 against 38 996.

[6] https://www.stockassessment.org/setStock.php?stock=whg.7b-ce-k_IBP_RefPts.

[7] https://community.ices.dk/ExpertGroups/benchmarks/2021/IBPCSWhiting/2021_Meeting_Documents/04.%20Working%20documents/WD03_WHGMSY_whg7bk_IBP_Sept2021_Ver2.html .

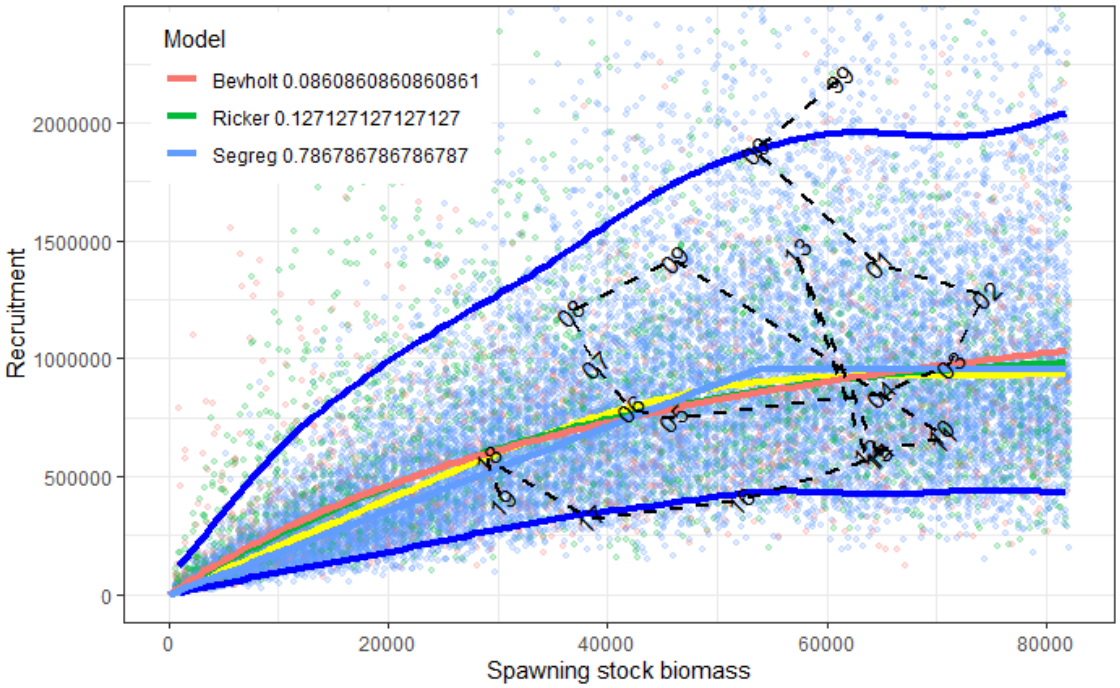


Figure 5.1. Stock–recruitment relations for Celtic Sea whiting red line: Beverton–Holt; green line: Ricker; blue line: segmented regression; yellow line represents the best fit over the three models).

The available time-series was truncated by removing the last two data years (2018-2019), as these are considered unreliable for reference point estimation. Bloss remained the same observation point (2008) as per WKCELTIC 2020. B_{lim} was set as B_{loss} = 36 571 t as per WKCELTIC 2020 and the reference points calculated using the segmented regression S_R relationship with a fixed breakpoint of $B_{lim} = B_{loss}$ (2008).

The revised reference points show a moderate upward revision of MSY $B_{trigger}$, B_{lim} and B_{pa} (Table 5.1). While F_{msy}, F_{lim} and F_{pa} have all reduced, although current guidelines mean $F_{pa} = F_{p05}$ in any event.

Table 5.1. Comparison of reference points from IBPCSWhiting 2021 vs. WKCELTIC 2020.

Reference Point	IBPCSWhiting 2021 Value	WKCELTIC 2020 Value	Rationale
MSY $B_{trigger}$	50 818 t	47 963 t	B_{pa}
F_{msy}	0.375	0.4	From EqSim with segmented regression and fixed breakpoint (Blim) capped to Fp0.5.
$F_{msyLower}$	0.315	0.332	Median lower point estimates of (F05)
$F_{msyUpper}$	0.375	0.4	Fp.05
Bl_{im}	36 571 t	34 516 t	B_{loss}; lowest observed SSB (2008) from which stock recovery was observed.
B_{pa}	50 818 t	47 963 t	B_{lim} combined with the assessment error; $B_{lim} \times \exp(1.645 \times \sigma)$; $\sigma = 0.20$ (default setting)
F_{lim}	0.64	0.89	F with 50% probability of SSB less than B_{lim}

Reference Point	IBPCSWhiting 2021 Value	WKCELTIC 2020 Value	Rationale
F_{pa}	0.375	0.4	Fp0.5; The F that leads to SSB ≥ B_{lim} with 95% probability.

6 Recommendations and future work needs

- The survey indices were truncated from 0–5 year olds down to 0–2 year olds as part of model fit optimization during the WKCELTIC 2020. This should be revisited again to ensure the model is not over fitting to the catch data.
- The fit of the French commercial LPUE index is not optimal and its use in the assessment should therefore be re-evaluated.
- Retrospective bias in the assessment shows an increase compared to the previous year's assessment, with Mohn's Rho values close to the acceptable limit, therefore this should be examined further.

7 References

ICES, 2021. Technical Guidelines - ICES fisheries management reference points for category 1 and 2 stocks (2021). https://doi.org/10.17895/ICES.ADVICE.7891.

ICES, 2020. Benchmark Workshop on Celtic Sea Stocks. https://doi.org/10.17895/ICES.PUB.5983.

Annex 1: List of participants

Name	Institute	Country (of institute)	E-mail
Hans Gerritsen	Marine Institute	Ireland	hans.Gerritsen@Marine.ie
Jesper Boje (reviewer)	National Institute of Aquatic Resources, The Technical University of Denmark (DTU Aqua)	Denmark	jbo@aqua.dtu.dk
Joanne Morgan (reviewer)	DFO	Canada	m.joanne.morgan@gmail.com
Rui Catarino	ICES Secretariat	Denmark	rui.catarino@ices.dk
Sara-Jane Moore	Marine Institute	Ireland	sara-jane.moore@marine.ie
Sofie Nimmegeers (chair)	Research Institute for agriculture, fisheries and food (ILVO)	Belgium	sofie.nimmegeers@ilvo.vlaanderen.be

Annex 2: Resolutions

IBPCSWhiting - Inter-Benchmark Protocol for the Celtic Seas whiting 2021

The Inter-Benchmark Protocol for Celtic Seas whiting 2021 that will serve as in Inter-Benchmark Protocol, Chaired by Sofie Nimmegeers, Belgium and reviewed by Joanne Morgan, Canada and Jesper Boje, Denmark will be established and meet online from the 23–24 September 2021 to:

c) Re-run the InterCatch raising procedure including the non-sampled component of landings and tune the model to the data.

d) Re-examine (and update if necessary) MSY and PA reference points according to ICES guidelines (see Technical document on reference points).

IBPCSWhiting will report by 30 September 2021 for the attention of ACOM and WGCSE.

Annex 3: Updating the Celtic Sea whiting assessment working document

David Stokes
The Marine Institute, Rinville, Oranmore, Co. Galway H91 R673
david.stoke@marine.ie

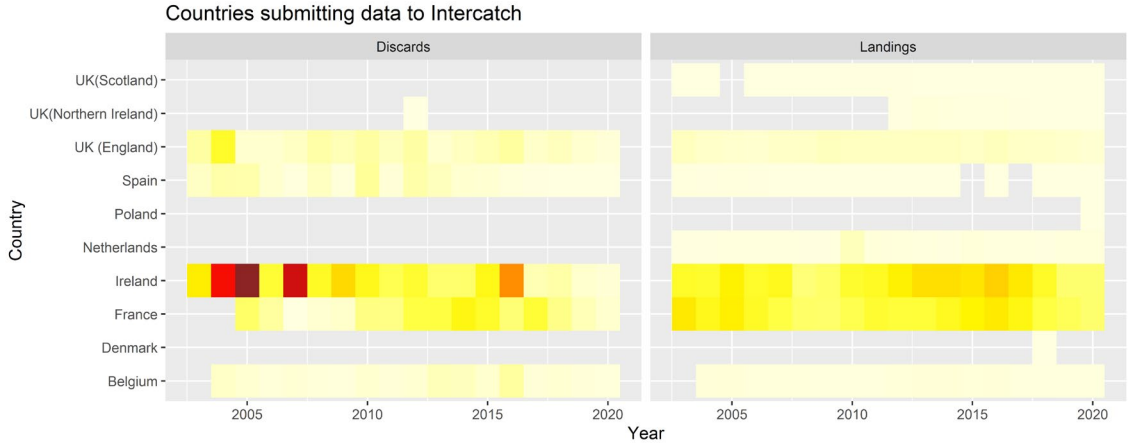
Introduction

The Celtic Sea whiting (whg.27.7b-ce-k) was benchmarked in 2020 along with the two other main gadoid stocks in this mixed fishery, cod and haddock. All input data back to 2003 was reviewed and the assessment was updated to State-Space stock assessment model SAM (ICES, 2020a). Reference points were also updated and the revised assessment presented and accepted at the 2020 assessment group meeting WGCSE2020 (ICES, 2020b). Subsequent to that, in 2021, an error in the shared code for the three stocks used to raise national catch and sampling data up to international catch-at-age data were identified just prior to WGCSE 2021. The code was developed as an R markdown document at WKCELTIC to allow consistent and transparent raising across the three stocks of national data submitted to Intercatch[8] to international catch-at-age data for the assessments.

Rectifying the issue caused an upward revision of SSB in particular that was more obvious in the early part of the time-series, but didn't appear too significant until reviewing the short-term forecast. Once the reference points were also revised using the updated data, the MSY advice suggested SSB would be below B_{pa} and B_{lim} in 2023. Given the significant change in perception of the stock since the benchmark it was agreed it was more appropriate to do a second detailed exploration of the data before providing advice later in the year. In this paper we present the review of the input data in the whg.27.7b-ce-k assessment.

Commercial catch data

A key term of reference for WKCELTIC was to develop a shared, transparent and standardized approach to exploration and raising of national catch data going into the three gadoid stock assessments (ICES, 2020a). This was addressed through an R markdown document used by the three stock coordinators to check and raise the full revised 2003-2020 time-series in Intercatch. Landings and discard data for whiting were reviewed and re-submitted by all 10 countries (Fig 1).



[8] https://www.ices.dk/data/data-portals/Pages/InterCatch.aspx

Fig 1. Revised time-series available in Intercatch for Celtic Sea whiting assessment.

For a generic approach to data exploration and for catch raising to be practical it was agreed the extensive list of métiers would be allocated to a higher level of Gear Groups: GNS_DEF, OTB_CRU, OTB_DEF, TBB_DEF and MIS_MIS. In so far as was possible, allocation of sampling to un-sampled métiers was also standardized across stocks using the same R markdown template and edited only if and where necessary. Sampled catches were available for four countries (Fig 2).

Fig 2. Number of length frequency samples per Country and Year, Quarter and Gear Group against the proportion of total catch in that metíers to the total annual catch across countries. Upper panel gives landings, lower panel shows discards.

Where catch data has no associated sampling, raising of catches to catch numbers-at-age is implemented using a simple hierarchy for available samples. Each iteration being progressively more of an 'average' sample until in the final step any catches that still are not matched to samples are simply raised by an aggregate sample for that year for that catch category (i.e. landings or discards).

Unsampled landings are paired with available samples in order of priority according to matching:

1) Country & Season & Year
2) Season & Year

3) Year

Unsampled discards are paired with available samples in order of priority according to matching:

1) Year, country and gear
2) Year and gear
3) Year

Where a table for catch was joined to a table for samples in this process, unfortunately the join had included Fleet1 (the higher level gear grouping) instead of Fleets (the original métiers code). This resulted in some lack of pairing samples with un-sampled catches in that raising process. The reduction in overall catch going into the assessment in 2020 compared to 2021 is more noticeable in the landings data and for the early part of the time-series (Fig 3).

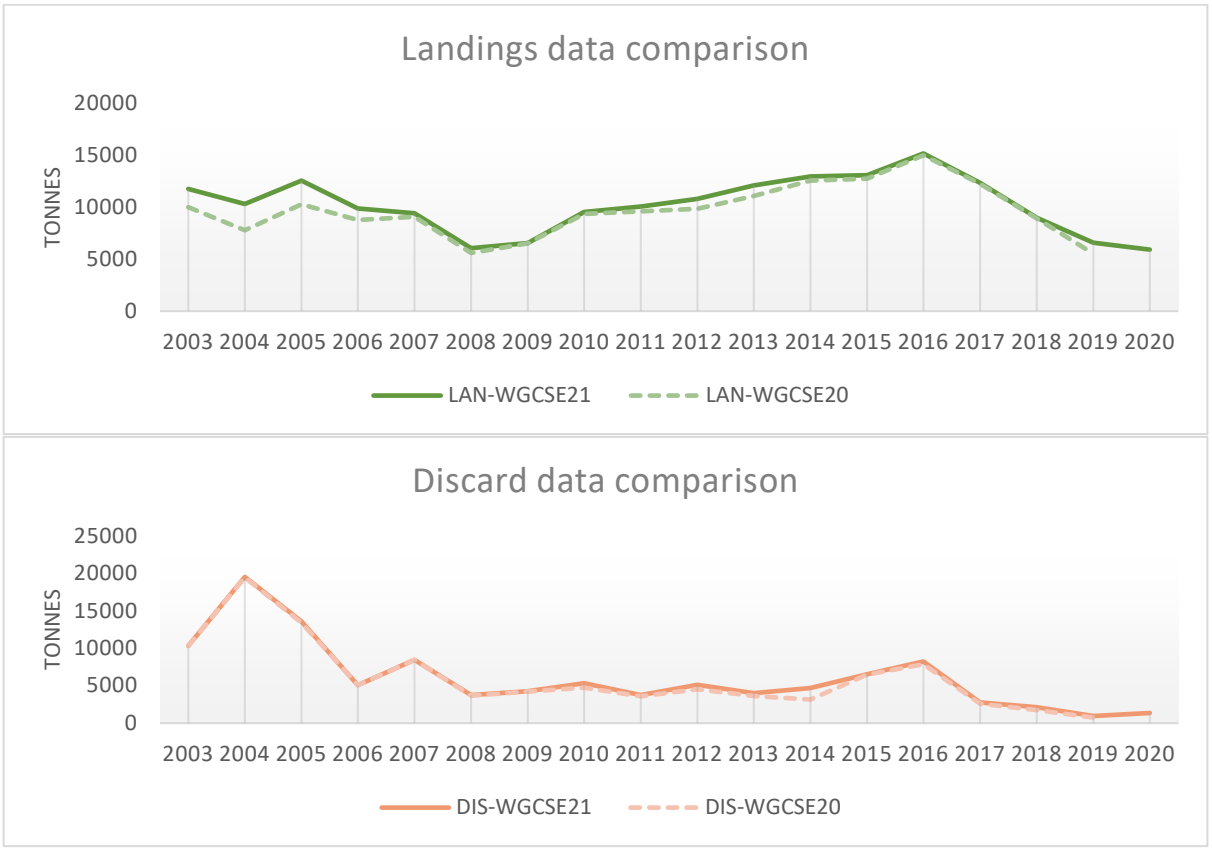


Fig 3. Comparison of raised international landings and discard data between assessment year 2020 and corrected data in assessment year 2021.

The resulting revision in the assessment of SBB in particular of that part in the time-series was significant, but matched the time frame of the data issue (Fig 4).

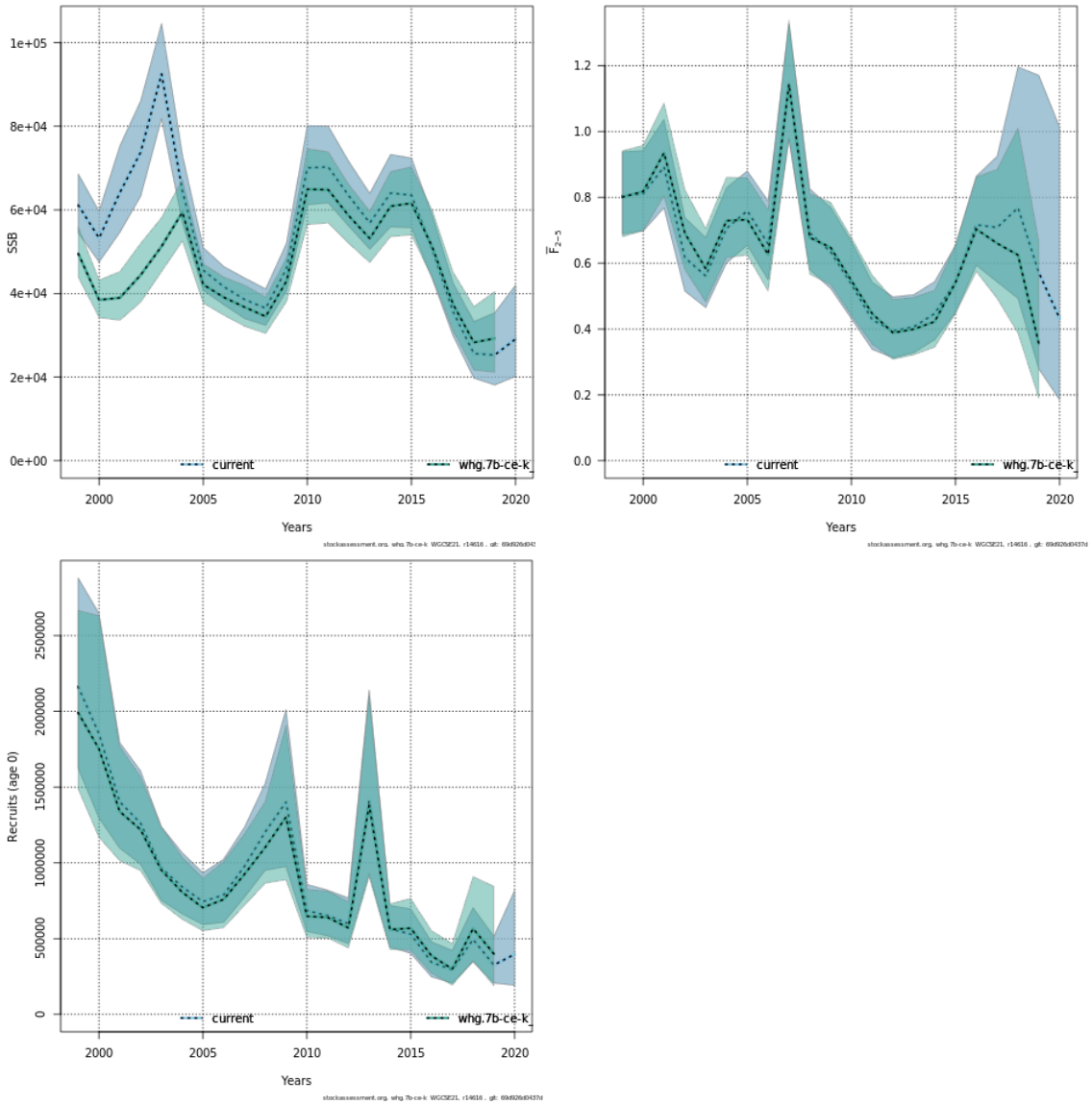


Fig 4. Summary plots from SAM assessment WGCSE 2021 (current – blue shading) compared with previous year WGCSE 2020 (green shaded confidence intervals). Greatest difference is early years of SSB (top left), with less noticeable differences in F (top right) and recruitment (bottom left).

In order to carry out a thorough review the raising procedure was run again in a new empty folder. A small amount of discard data for Belgium for data year 2019 (79t) and 2020 (65t) was identified as being available in accessions which was uploaded to Intercatch and imported again. The new files from Intercatch were checked and contained only the additional tons for Belgium as expected, there were no other edits to these or previous year's data.

Having raised the national data again to catch numbers-at-age, the assessment input files were created again using the same Excel workbook as set up during the benchmark. Essentially the canum_summary.csv file output from the R markdown[9] is imported and once the pivot tables refresh the formatted input files for the assessment can be copied directly out.

[9] https://community.ices.dk/ExpertGroups/benchmarks/2021/IBPCSWhiting/2021_Meeting_Documents/04.%20Working%20documents/WD02_aggregate_IC_data_whg.27.7b-ce-k_Sept_2021.html

The final SAM assessment in stockassessment.org from WGCSE2021 was cloned and saved (https://www.stockassessment.org/setStock.php?stock=whg.7b-ce-k_IBP) and all input files imported again. On running the assessment however, the revised high peak of SSB in 2003 seen at WGCSE2021 was not as evident (Fig 5) so further checking of the input data were carried out.

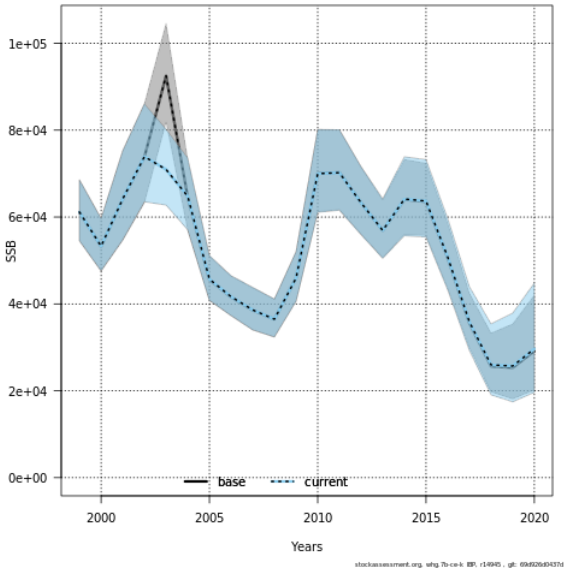


Fig 5. Plot of SSB from SAM assessment for Celtic Sea whiting IBP (current - blue shaded CI) against the recent WGCSE 2021 assessment (base – grey shaded CI).

Again all the input files were checked against those used for WGCSE21. All matched except an anomaly identified in the stock weights file. This is created by taking the mean weights at age and passing it to a Rivard correction utility[10] from the NOAA toolbox. The calculator converts mean weights at age to weights at age for Jan 1st. This is the only file that requires moving data between applications and not automatically updated and formatted within the workbook. A difference was discovered in the stock weight ssyued in WHCSE2021 and WKCELTIC 2020.

Stock weights

The anomaly with the stock weights was unusual for a copy and paste error as it did not affect all rows, only years 1999-2003 (Fig 6). Mean weights at age data prior to 2003 was not revised in the benchmark and remains fixed and stock weights for Years > 2002 therefore should not change. The cause seems to be that the mean weights at age table in the workbook was extended back to 1999 to simplify using the Rivard calculator for the full dataset. Unfortunately, stock weights were used instead of mean weights at age for the 1999-2002 data and therefore were corrected a second time when passed into the Rivard correction utility as part of the full dataset.

Comparing the revised stock weights file with that from WGCSE2020 shows the discrepancy with 1999-2002 data (Fig 6). Smaller differences in more recent years reflect the catch data revision discussed above, but the key result being the underestimate of SSB at that point. Comparison of the revised stock weights file with WGCSE21 shows smaller differences and ostensibly around the years of revised catch data so values are consistent with WGCSE2021. The obvious anomaly being ages 1-6 for 2003, but additional checks are in place to highlight any future change in these values specifically given it is the only data table that undergoes processing outside the R markdown process and resulting workbook.

[10] RivardCalc 2.0 https://nmfs-fish-tools.github.io/RIVARD/

	0	1	2	3	4	5	6	7
1999	1.589	1.288	1.336	1.217	1.148	1.148	1.104	1.000
2000	1.876	1.594	1.287	1.336	1.217	1.148	1.148	1.000
2001	1.748	1.885	1.618	1.298	1.243	1.167	1.155	1.000
2002	3.200	1.749	1.582	1.501	1.246	1.156	1.130	1.000
2003	0.984	1.472	1.333	1.165	1.180	1.089	1.074	1.001
2004	0.989	1.016	1.014	1.023	1.030	1.013	1.007	0.981
2005	0.995	1.010	1.022	1.010	1.006	1.019	0.964	1.017
2006	1.000	1.006	1.014	1.014	0.996	0.989	1.007	1.002
2007	0.989	1.000	1.008	1.008	1.008	0.999	1.000	1.005
2008	1.000	1.007	1.003	1.002	1.001	1.003	1.004	0.995
2009	0.987	1.001	1.009	1.003	1.000	0.999	1.002	1.000
2010	1.015	1.010	1.002	1.001	1.000	0.999	1.003	1.012
2011	1.042	1.016	1.017	1.006	1.001	1.004	1.002	1.004
2012	0.997	1.034	1.005	1.011	1.016	1.009	0.995	0.995
2013	0.981	1.010	1.024	1.009	1.011	1.018	1.016	1.030
2014	1.117	1.026	1.029	1.012	1.007	0.999	1.003	1.001
2015	1.003	1.039	1.023	1.024	1.005	0.994	0.994	1.002
2016	1.054	0.995	0.994	0.998	1.001	1.002	0.999	0.998
2017	1.038	1.022	0.992	0.990	0.996	0.999	1.000	0.999
2018	0.940	0.997	0.990	0.989	0.993	0.997	1.000	1.000
2019	1.049	1.050	0.997	0.992	0.991	0.993	0.996	0.985
2020	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!

	0	1	2	3	4	5	6	7
1999	1.006	0.999	1.000	1.000	1.000	1.000	1.000	1.000
2000	1.006	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2001	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2002	0.877	1.000	0.999	1.000	1.000	1.000	1.000	1.000
2003	1.000	0.570	0.798	0.802	0.917	0.854	1.161	1.000
2004	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2005	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2006	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2007	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2008	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2009	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2010	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2011	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2012	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2013	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2014	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2015	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2016	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2017	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2018	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2019	0.992	0.999	0.999	1.000	1.000	1.000	1.000	1.000
2020	1.000	1.006	1.004	1.001	1.001	1.001	1.000	0.998

Fig 6. Check of Rivard corrected stock weights input files. Corrected file for IBP is divided by the stock weights file used at WGCSE2020 (left) and WGCSE2021 file (right). A significant shift in values can be seen in the early part of the time-series only. Slight differences in 2019-2020 are attributed to the updated discard data from Belgium.

With the corrected data, the SAM assessment summary plots show a rescaling of SBB in the early part of the time-series with a peak in 2002. Fishing mortality and recruitment through the time-series remain largely unaffected by this change.

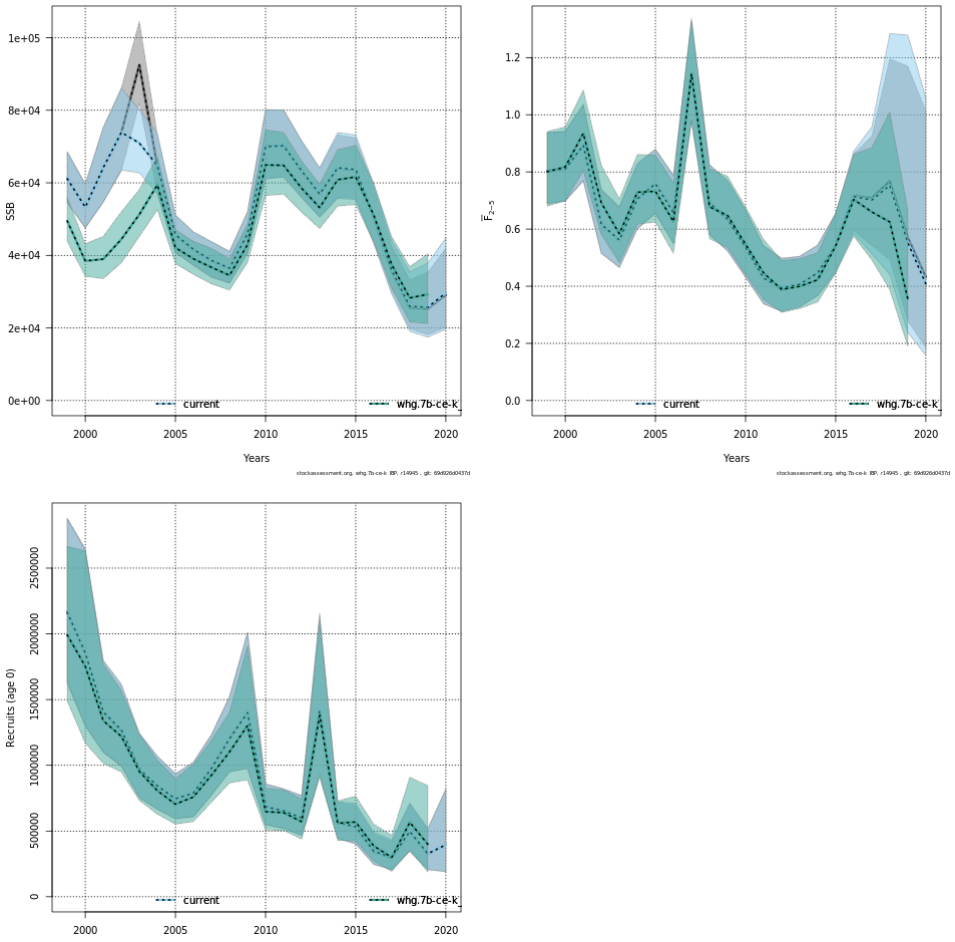


Fig 7. Summary plots from SAM assessment IBP2021 (current – blue shading) compared with previous year WGCSE 2020 (green shaded confidence intervals). The WGCSE2021 is included for comparison with grey shaded confidence inetrvals. Greatest difference is early years of SSB (top left), with less noticeable differences in F (top right) and recruitment (bottom left).

Looking at the SSB estimates historically, the revised data are more in line also with the pre-benchmark XSA assessment carried out up to, and including, WGCSE 2019[11] (Fig 8).

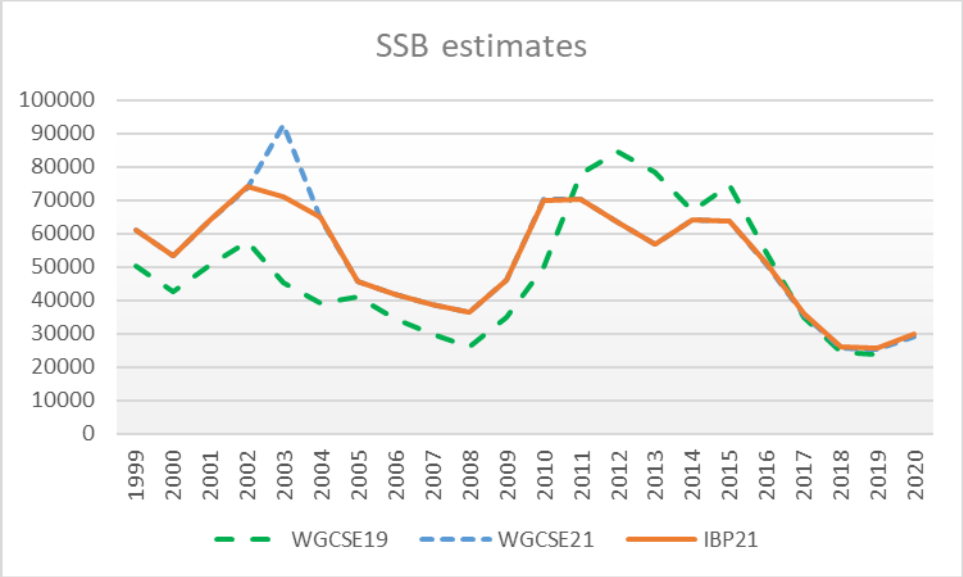


Fig 8. SSB estimates from the WGCSE2019 XSA assessment compared to SAM assessment WGGCSE21 and corrected IBP2021.

SAM assessment

The resulting SAM assessment[12] is summarized below showing the outcome using revised data as above with only the current run for clarity. SSB shows a clear peak early in the time-series in 2002 corresponding to the peak of recruitment 3 years prior in 1999. Later recruitment peaks in 2009 and 2013 seem to be followed increased SSB 2 years later. Overall recruitment is noisy, but has been on a downward trend. Likewise, SSB has been variable, but appears to be growing slowly in the last 2 years. Fishing mortality has been progressively dropping since 2018, but model uncertainty remains high since that 2018 inflection point (Fig 9).

[11] https://community.ices.dk/Advice/Advice2019/CelticSeas.

[12] https://www.stockassessment.org/setStock.php?stock=whg.7b-ce-k_IBP_final.

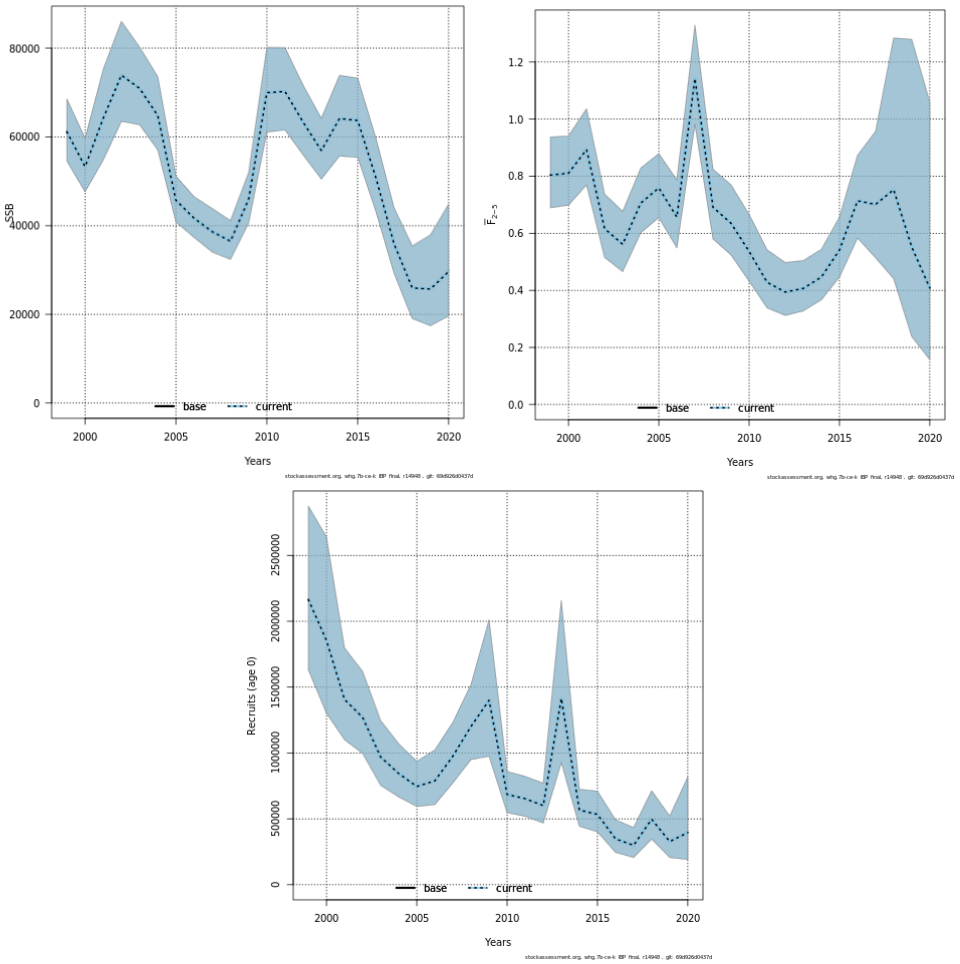


Fig 9. SSB, Fishing mortality (F) and Recruitment estimates from the IBP21 SAM assessment.

Model fits are generally good for the catch data with survey observations often higher for younger ages and conversely lower than model predictions for older fish (Fig 10). Model fit to the French commercial biomass index is not optimal, but was evaluated at benchmark and also improves Mohn's Rho estimates so was recommended for inclusion in the Celtic Sea whiting assessment.

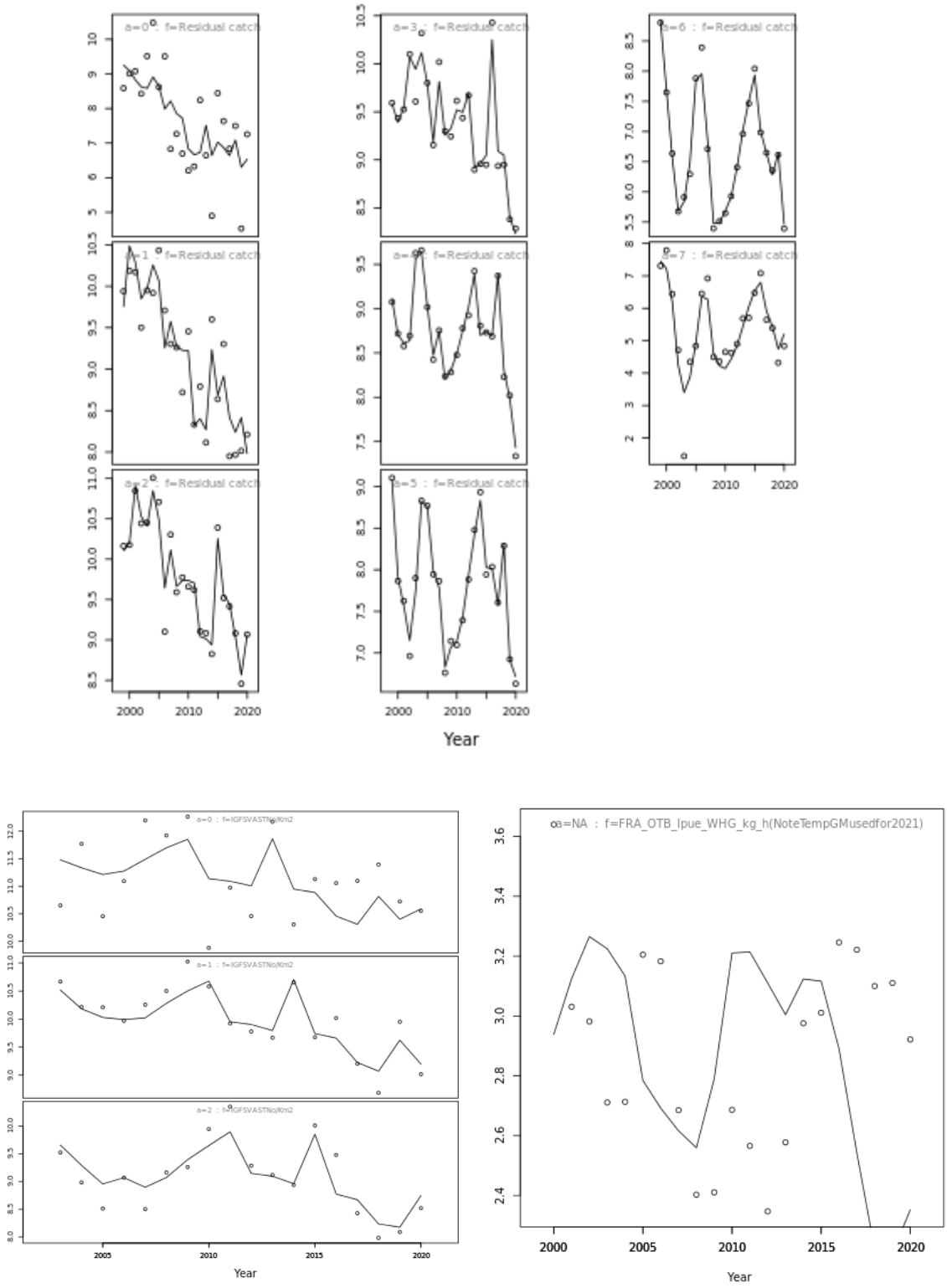


Fig 10. SAM model predictions (solid line) against observations for catch data (top left: age0-7), survey data (bottom left: 0-2) and French LPUE index (bottom right).

Standardized 'one ahead' residuals (Fig 11) show some reasonable cohort tracking with the VAST survey index in particular tracking the stronger 2008, 2009 and 2013 recruitment well. The catch data picks up the 2013 year class as age 1 the following year and tracks positively for next three years.

Fig 11. Standardized residuals from SAM by year class plotted for time-series. Catch data (top panel), French commercial LPUE index (middle panel) and VAST survey index (bottom panel).

Retrospective bias in the assessment shows a fairly consistent downward revision in SSB and upward revision in F, but with Mohn's Rho values toward the high end of acceptable (Fig 12).

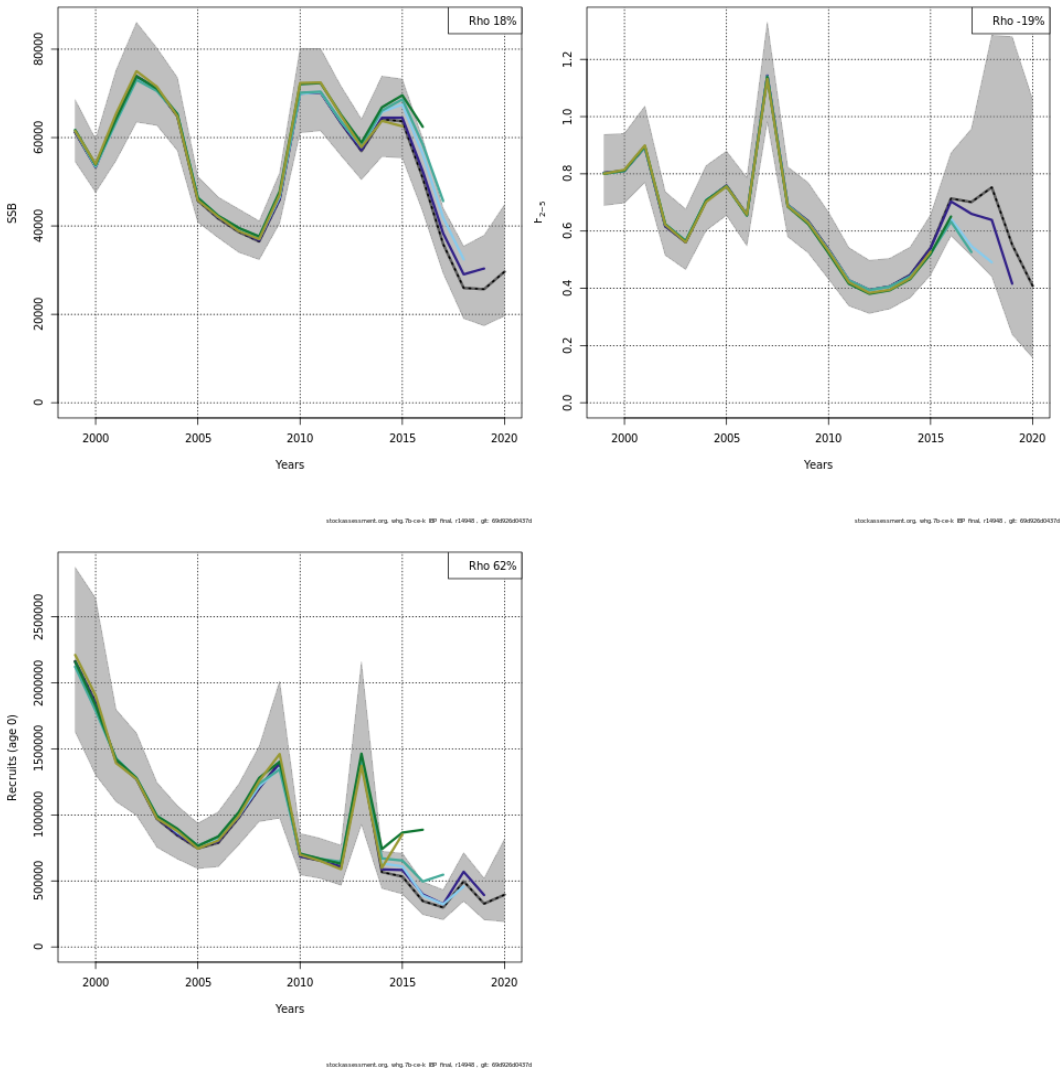


Fig 12. Retrospective runs for Celtic Sea whiting giving Mohn's Rho values for bias.

References

ICES, 2020a. Benchmark Workshop on Celtic Sea Stocks. https://doi.org/10.17895/ICES.PUB.5983.

ICES, 2020b. Working Group for the Celtic Seas Ecoregion (WGCSE). https://doi.org/10.17895/ices.pub.5978.

Annex 4: Reviewer reports

IBPCSWhiting - Inter-Benchmark Protocol for the Celtic Seas whiting 2021

Reviewer's Report Joanne Morgan, Fisheries and Oceans Canada

This Inter-Benchmark was called to:

a) Re-run the InterCatch raising procedure including the non-sampled component of landings and tune the model to the data.
b) Re-examine (and update if necessary) MSY and PA reference points according to ICES guidelines (see Technical document on reference points).

The main factors considered in this IBP were the correction to the catch-at-age and its impact on the new assessment model and the subsequent impact on reference points.

The catch-at-age issue had three components: an error that was detected prior to WGCSE 2021 in the R program to do the calculation, the inclusion of some new data, and a change in the calculated stock weights for early years of the time-series.

The error in the R program was discussed extensively during the IBP and seems to have been fixed. The program and its resulting markdown document should make things more transparent and reproducible in future. The extra data seem to be relevant and should be included. One of the major impacts on the SSB seems to have been the change in the stock weights at age. This was also discussed extensively and checked during the IBP. The current data seem to be correct and the error seems to be with previously used data.

The impact of these changes relative to the previous year (WGSCE 2020) is mainly on the SSB in 2003 which with the current data are lower and forms less of a peak. This seems more realistic than the large peak in that year in the previous assessment. This is mainly the result of the adjustment in the stock weights.

These issues (the error in the catch numbers-at-age and the stock weights) are not easy to describe but are the key to the IBP. There needs to be great care given to adequately describing them in the final report.

There are some retrospective patterns with SSB being revised down and F revised upwards with each additional year. This could be a problem and will need to be carefully examined in each assessment. Mohn's Rho for these are somewhat high and very high for recruitment. This seems to be mainly due to the large difference in the first and second year of the retrospective runs compared to the last several years. Such jumps in estimates with the addition of one year of data can indicate instability in the model and this needs to be closely monitored.

The analyses of reference points seem logical. It is not clear-cut whether the type of S/R should be category 3 or category 5 but 5 was chosen and the rational is sound. This does need to be clearly explained and described in the report. This results in Bloss as Blim. The simulations to determine other reference points requires a stock recruit relationship and based on fitting of several models a segmented regression was chosen.

Several small issues with captioning and text were given to the author for revision.

Conclusions/recommendations

The changes to the catch-at-age and stock weights, as well as the reference point calculations are logical and acceptable.

The error in the catch numbers-at-age and the stock weights are not easy to describe but are the key to the IBP. There needs to be great care given to adequately describing them in the final report.

There are some retrospective patterns with SSB being revised down and F revised upwards with each additional year. This could be a problem and will need to be carefully examined in each assessment. Mohn's Rho for these are somewhat high and very high for recruitment. This seems to be mainly due to the large difference in the first and second year of the retrospective runs compared to the last several years. Such jumps in estimates with the addition of one year of data can indicate instability in the model and needs to be closely monitored.

The choice of category 5 S/R for the setting of Blim should be clearly explained in the report.

It would be helpful if a description of the reason for the IBP were included in the ToR.

IBPCSWhiting - Inter-Benchmark Protocol for the Celtic Seas whiting 2021

Reviewer's Report Jesper Boje, DTU Aqua, Denmark

This Inter-Benchmark was called to:

a) Re-run the InterCatch raising procedure including the non-sampled component of landings and tune the model to the data.
b) Re-examine (and update if necessary) MSY and PA reference points according to ICES guidelines (see Technical document on reference points).

When the issues for ToR appeared was somewhat unclear prior to the IBP, i.e. what issues was discovered after benchmark 2020 at WGCSE in 2020 and 2021 and after 2021 meeting. This narrative should be included in the IPB report in order to through light on the decisions from the various meeting and the process of solving them.

Given the ToR for this IBP, generic issues on the assessment not affected by InterCatch re-run and re-estimation of reference points, are only raised here and could be added to the to-do list for the next benchmark.

Issues with the pre-InterCatch procedure for normal InterCatch use and the InterCatch allocation of unsampled catches was ToR a. Both procedures was conducted in R and thus not within the online environment offered by ICES. My skills in R is poor and I was therefore not in a position to review the R code. However, a rectified code issued as a markdown document was agreed upon. The corrected CANUM and WEST input to the SAM assessment did led to minor recent changes in SSB and F but had an effect on SSB in especially 2003 due to WEST corrections for that year.

The ranking of the grouping variables was discussed on the raising procedure; presently nation is ranked higher than both gear and area, i.e. nation is assumed to better reflect length/age distribution for unsampled catches that gear and area. This is apparently the default procedure and not part of the present ToR. However, the WGCS might consider this issue more in depth for a future benchmark.

Apart from changes in SSB in the early part of the time-series (2000-2005) the perception of stock and fishery remains unchanged. However, for future benchmarks some issue are worth mentioning. A biased retro pattern is seen for SSB and F with overestimation and underestimation, respectively. Mohn's rho for both are just within the acceptable range and the past estimates are all within the confidence limits of the present estimates. However, the pattern is worrying and should be examined in future if it continues.

The commercial fleet with an aggregated SSB index do not seem to follow neither catch data nor the survey indices. Large blocks of negative and positive residuals are obvious from the present assessment. A run without this tuning fleet improved the assessment (AIC) and gave more precise SSB and F estimates (less uncertainty), but slightly increased the Mohn's rho for the retrospective plots. So dependent on the main objective on the assessment quality, this might be a generic issue for ACOM to reflect on.

Given the changes in InterCatch, though minor, reference points was agreed to be recalculated. A discussion evolved on the data points used for the calculation. The recent recruitment estimates (2018 and 2019) are considered associated with high uncertainty and therefore omitted from the stock–recruitment relationship in the EqSim calculations. These two observations were therefore not considered when estimating Blim (here lowest observed SSB) and when estimation the relationship for use (segmented regression). The recommended procedure for reference

points in EqSim was followed. The estimated Fmsy (0.53) was capped by the Fp05 at 0.319. Overall the reference points changed only slightly from the past (2020) definitions.

Summary and recommendations

The issues with sample aggregation and allocation in Intercatch was solved and WEST issues for the early period likewise. The markdown documents are valuable for examining and documenting the procedures.

The re-calculation of reference points was done appropriately and the new reference points adopted.

The two ToRs was thus addressed adequately.

For future benchmarks the WGCSE should consider the use of the commercial fleet. Also the retrospective behaviour of SSB and F should be examined further.

ACOM should provide guidance/elaboration on quality improvement of assessments in the case where AIC and confidence limits improves but retrospective behaviour get worse at the same time.