WORKSHOP ON EVALUATING SURVEY INFORMATION CELTIC SEA GADOIDS (WKESIG)

VOLUME 2 | ISSUE 107

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 I © 2020 International Council for the Exploration of the Sea

ICES Scientific Reports

Volume 2 | Issue 107

ICES WORKSHOP ON EVALUATING SURVEY INFORMATION CELTIC SEA GADOIDS (WKESIG)

Recommended format for purpose of citation:

ICES. 2020. ICES Workshop on evaluating survey information Celtic Sea gadoids (WKESIG).
ICES Scientific Reports. 2:107. 26 pp. http://doi.org/10.17895/ices.pub.7574

Editors

David Stokes

Authors

Olav Nikolai Breivik • Paul Dolder • Hans Gerritsen • Natoya Jourdain • Cóilín Minto • Lionel Pawlowski • Sven Kupchus • Marianne Robert • David Stokes

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

Contents

i Executive summary

The Workshop on evaluating survey information on Celtic Sea gadoids (WKESIG) provides expertise for the survey indices that contribute to the Benchmark Workshop on Celtic Sea Stocks (WKCELTIC). Three of the largest stocks assessed by the Working Group for the Celtic Seas Ecoregion (WGCSE), namely Cod 7e_k, Haddock 7b_k and Whiting 7b_k, form part of a significant mixed demersal fishery in the Celtic Sea. The assessments rely heavily on survey indices, which in all cases use combined survey indices between the Irish Ground Fish Surveys (IE-IGFS) and the French Surveys (EVHOE).

A number of survey data issues were highlighted for potential review by the Working Group on Improving Use of Survey Data for Assessment and Advice (WGISDAA). Key issues include: (1) how survey data that vary across indices are standardized and combined, (2) unavoidable gaps in survey coverage that have occurred in recent years, and (3) estimates of uncertainty that are not routinely calculated for these index calculations.

Having explored variability of survey effort, it was agreed data gaps would be addressed to allow production of indices by swept area for the benchmark. Additional data such as speed, trawl settling time, and sea conditions were discussed as useful measures of observation error. A few modern modelling approaches were also discussed and should be explored further.

Migration patterns should be included as supplementary information in the benchmark, but it was not clear how this information could be integrated into assessments yet in a quantitative way.

Modelling of average length keys (ALK) data highlighted the precision gains achievable when spatial variability can be accounted for. This ongoing work is a progression of the methods currently used for ALK 'fill-ins' for current indices, so the working group will continue to follow its development.

Dealing with overall data gaps was addressed by presentation of a case study on Celtic Sea whiting index calculation using a vector autoregressive spatio-temporal model (VAST). It was shown that the model had good stability and predictive accuracy even when large amounts of input data were removed from a given year, to simulate vessel breakdown for example. However, the ability to model the missing values was heavily influenced by how 'average' a survey year was. In unusually high or low abundance years the capacity to 'model your way out' of data gaps is reduced significantly. Overall the modelling approach offered potential to address spatial variability of the data, estimate uncertainty about observation and process error separately and also reduce the impact of unavoidable data gaps.

.

ii Expert group information

Expert group name	Workshop for Evaluating Survey Information on Celtic Sea Gadoids (WKESIG)
Expert group cycle	Annual
Year cycle started	2019
Reporting year in cycle	1/1
Chair(s)	David Stokes, Ireland
Meeting venue(s) and dates	4th – 5th February 2019, Galway, Ireland (9 participants)

1 Terms of reference (ToR)

WKESIG – Workshop on evaluating survey information Celtic Sea gadoids

2017/2/EOSG08 The **Workshop on evaluating survey information for Celtic Sea gadoids (WKESIG)**, chaired by David Stokes, Ireland, will meet in 4-5 February 2019 at the at the Marine Institute, Galway, Ireland to:

a) Review and consider the quality and availability of survey data going into the assessment of cod, haddock and whiting as requested by WGCSE2017 (Science plan codes 5.1);

b) Evaluate the potential to improve current survey indices by use of additional information such as standardizing by swept area or using model based index approaches (Science plan codes 5.1);

c) Review and standardize methods for evaluating and constructing indices including applying and filling in ALKs, estimating uncertainty and, where desirable, combining or complementing surveys with other data sources (Science plan codes 5.1).

2 Introduction

Three of the largest stocks assessed by the Working Group for the Celtic Seas Ecoregion (WGCSE), namely Cod 7e_k, Haddock 7b_k and Whiting 7b_k, form part of a significant mixed demersal fishery in the Celtic Sea. These assessments rely heavily on survey indices which in all cases use at least one combined survey index between Ireland (IE-IGFS) and France (EVHOE). How these data are standardized and combined is somewhat different across stocks and achieved by R code passed through earlier Benchmarks, but not currently published or documented in detail.

These stocks are scheduled to be benchmarked together as part of WKCELTIC, starting February 2019. Given a number of survey specific data questions were identified by WGCSE the Working Group on Improving use of Survey Data for Assessment and Advice (WGISDAA) was asked to facilitate the current workshop. The objective being to review methods currently used and available alternatives for survey index preparation in advance of the benchmark process.

There was overlap across ToRs in some of the presentations so for simplicity the report is structured around topics rather than ToRs.

3 Availability and quality of survey data.

3.1 History

One presentation was given under Tor A summarizing the current status of survey data inputs for Celtic Sea cod, haddock and whiting assessments.

Currently two IBTS coordinated surveys are available for the Celtic Sea area and both used within the assessments for all three stocks (Fig 3.1.1). The French Southern Atlantic Bottom Trawl Survey (EVHOE 1997-present) and the Irish Groundfish Survey (IGFS 2003-present).

Historically indices were used independently by survey, but often conflicted in terms of cohort tracking in cod for example[1]. No analytical assessment was performed for Cod between 2008 and 2009 and only exploratory assessments in 2010 – 2011. It was felt likely, for cod in particular, that a stronger juvenile signal was coming from the increased number of coastal IGFS stations. In contrast, ages classes tended to be of older fish in the more offshore central Celtic Sea stations of the EVHOE survey.

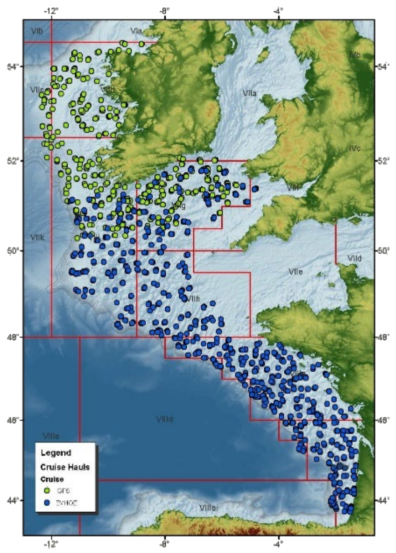


Figure 3-1.1 Survey station positions for the FR-EVHOE and Celtic Sea portion of the IE-IGFS surveys.

Production of combined indices for cod and haddock for WKROUND2012 was done to address signal to noise ratio issues for cod in particular. More generally it also provided single indices with almost complete coverage of the assessment area. Two approaches were taken however

i. Cod data were combined using a spatial grid (Fig 3.1.2) where average CPUE per grid cell is then averaged to produce an annual mean CPUE in numbers per age per hour (WKROUND 2012: WD 10). The spatial grid helps account for slight spatial differences in survey effort between surveys.

---

[1] ICES. 2012. Report of the Benchmark Workshop on Western Waters Roundfish (WKROUND), 22–29 February 2012, Aberdeen, UK. ICES CM 2012/ACOM:49. 283 pp.

ii. Relative catch at length for Haddock between surveys was checked for neighbouring hauls using linear modelling. Length frequencies were shown to be highly correlated (R^2 = 0.71) and the data then combined as a geometric mean number-at-age per hour. However, the contribution of each survey to the combined index in this case is weighted by the spatial coverage of each survey in the Celtic Sea (WKROUND 2012: WD 12).

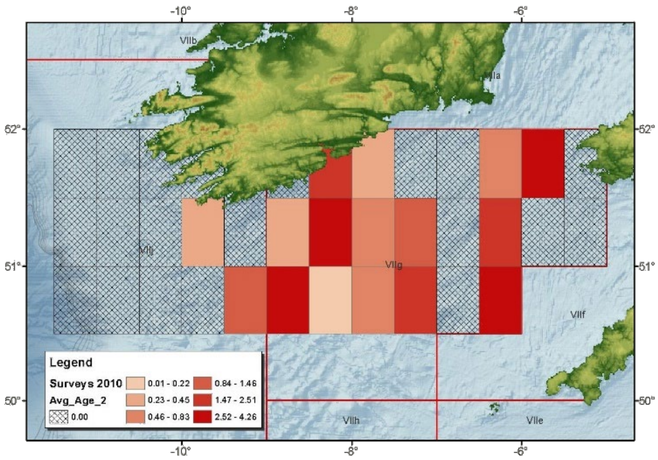


Figure 3.1.2 Grid used to combine FR-EVHOE and IE-IGFS cod survey data as a combined index. Example shown is for Age 2 cod in 2010 in catch number per Hr.

Following similar logic and procedures as that for Celtic Sea cod, a combined index for whiting was provided and approved during the relevant benchmark WKCELT2012[2]. An extended grid was used to include areas where historically whiting occurred within the surveys over multiple years (Fig 3.1.3).

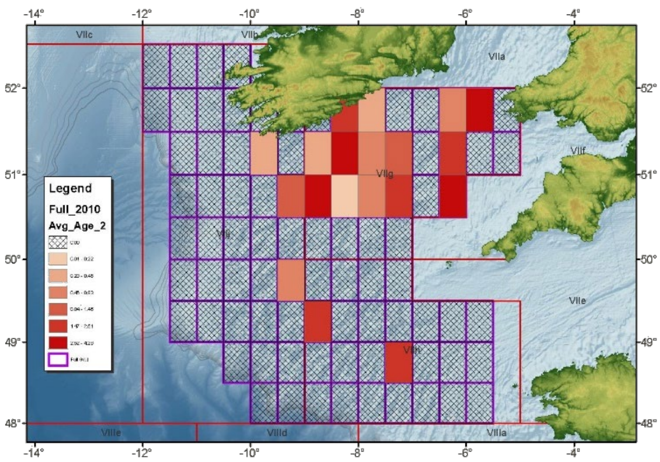


Figure 3.1.3 Grid used to combine FR-EVHOE and IE-IGFS whiting survey data as a combined index. Example shown is for Age 2 whiting in 2010 in catch number per Hr

[2] ICES. 2014. Report of the Benchmark Workshop on Celtic Sea stocks (WKCELT), 3–7 February 2014, ICES Headquarters, Copenhagen, Denmark. ICES CM 2014\ACOM:42. 194 pp.

All three indices apply within survey age–length data, stratified where possible, on the length frequency by haul to produce the catch number-at-age per haul. Where an age is missing a multinomial regression, model is used to "fill-in" missing ages[3].

3.2 Current issues

The EVHOE and IGFS are part of the IBTS Q4 standardized and coordinated surveys, use the same GOV Type-A trawl and have done comparative calibration work historically. Spatial extent is good relative to assessment area, commercial activity and known trawlable ground. The key issues then in relation to this workshop were i) missing data; ii) estimating uncertainty.

i. Full analytical assessments are now established for these stocks, but a number of unavoidable issues in recent years, such as vessel break down, have resulted in a significant reduction in survey data for the given year. The methods to produce combined indices currently cannot account for missing data and may therefore lead to an increased bias in the estimate for that year. If data is not collected in an area of generally larger fish offshore due to weather for example, the estimate will suggest proportionately smaller fish in the population. The limited IBTS survey overlap in the NE Atlantic compared to the North Sea and the reliance on survey data in these assessments was highlighted and the workshop discussed the merits of modelled indices (see Tor b).
ii. The second topic discussed was sources of uncertainty in survey indices generally and the availability of data to evaluate and reduce that. Much survey coordination effort goes into standardization across sampling procedures to ensure differences in relative catch reflect variation in the population and is not confounded with observation error in sampling the fish[4]. This important assumption around standardized surveys (survey catchability = 1), is the combination of both the observation and process error. Standardization can achieve much about reducing observation error, however aspects such as human error, weather conditions, sea temperature and so forth are largely beyond our control. Fortunately they are often measured precisely and routinely by surveys so, could these measurements be reviewed to distil out survey observation error separately to population variability?

As an illustration of weather conditions during the IGFS, box plots of windspeed and vessel heave, pitch and roll over a number of years were presented. The assumption being vessel movement at the surface is likely to affect trawl contact at the seabed, when corrected for depth.

[3] Hans D. Gerritsen, David McGrath, and Colm Lordan, 'A Simple Method for Comparing Age–Length Keys Reveals Significant Regional Differences within a Single Stock of Haddock (Melanogrammus Aeglefinus)', *ICES Journal of Marine Science: Journal Du Conseil* 63, no. 6 (1 January 2006): 1096–1100, https://doi.org/10.1016/j.icesjms.2006.04.008.

[4] Ahrestani, F.S., Hebblewhite, M. & Post, E. The importance of observation versus process error in analyses of global ungulate populations. Sci. Rep. 3, 3125; DOI:10.1038/srep03125 (2013).

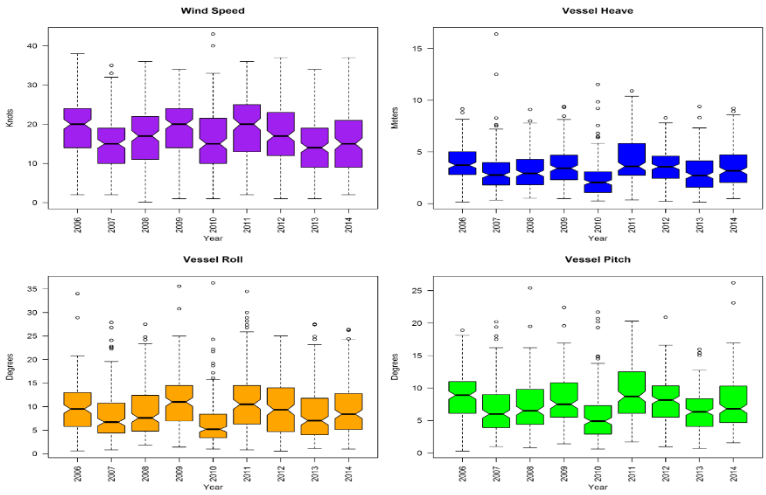


Figure 3.2.1 Boxplots of true windspeed, vessel heave, pitch and roll encountered by the *R.V. Celtic Explorer* on the IGFS during valid fishing tows.

It is worth noting that 2010 was a particularly poor weather year and at a certain point any vessel will need to orientate into the weather to fish, or indeed abandon survey work altogether. In these circumstances' vessel roll can appear to reduce with swell being on the bow during fishing rather than beam on. However, pitching of the vessel will increase, introducing more 'whipping' energy being passed down the warps, which has the potential to impact on ground contact depending on depth. Worst-case scenario the vessel has to shelter and cease fishing until weather abates. At a certain point data for weather parameters may even suggest improved sampling conditions (for the achieved tows) even while prevailing conditions are generally poor. The overall impact might then only be seen as a reduction in total effort for that survey such as total swept area (Fig 3.2.2).

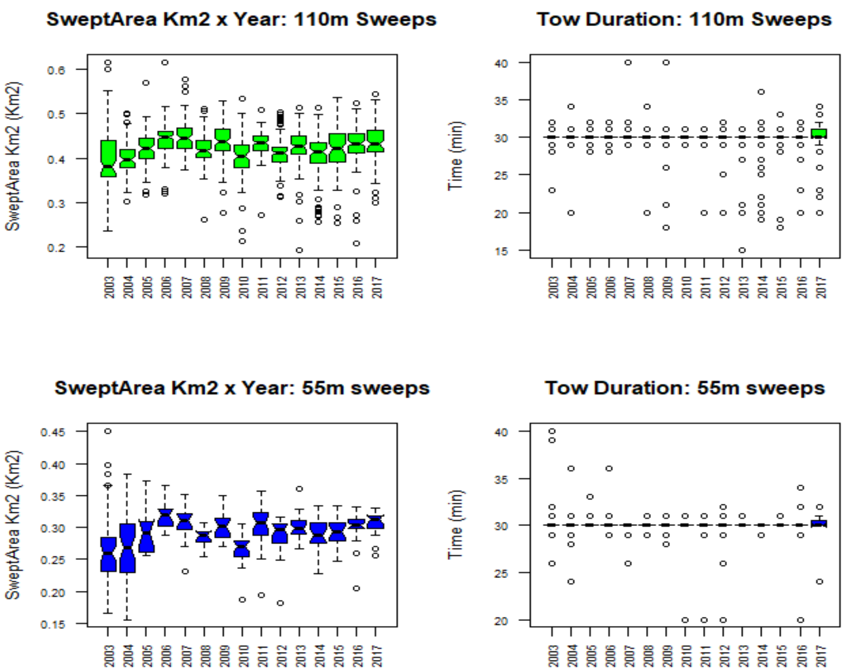


Figure 3.2.2 Boxplots of swept area (left panels) and tow duration (right panels) for the IGFS 2003-2017. Upper panels are for the IBTS longer 110m sweeps configuration in depths > 80m. Other than the first year in the time-series (2003) the poorest survey coverage was 2010.

While the initial year of the survey was followed by a few inevitable changes, the time-series shows particularly poor coverage in 2010 in terms of swept area while time, the unit of effort actually used for the index, remains relatively unaffected over the years. It is also obviously a relatively constant 30min between short or long sweeps despite the area sampled in those 30min being considerably different. We can see clear differences then between the extent of area sampled between years, the extent of deep vs. shallower areas within a year and the conditions that may have influenced catchability. Standardizing CPUE to 30min will evidently miss this information routinely collected by the survey.

A quick comparison with the second survey in the index (EVHOE) shows no such issues in 2010 for example (Fig 3.2.3) and possibly an annual weighting of survey data could be considered. Currently not all gear parameter data is readily available in DATRAS for many surveys and quite a few points had to be estimated for EVHOE based on formula agreed at IBTS for DATRAS products. This makes standardizing even by swept area sometimes problematic. However, there is ongoing work at IBTS to QC and improve these datasets as well as discussion directly between the IGFS and EVHOE participants at the workshop.

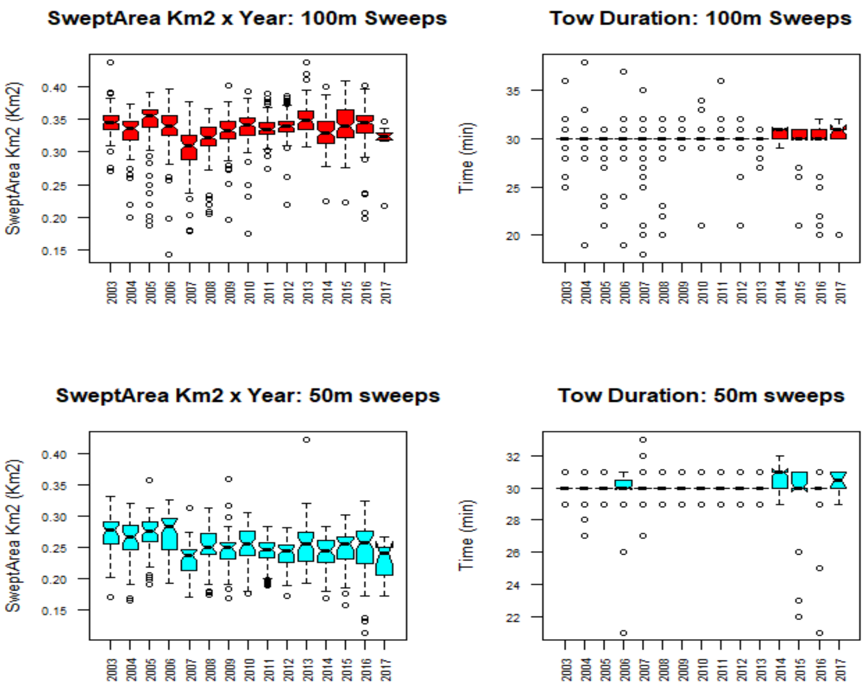


Figure 3.2.3 Boxplots of swept area (left panels) and tow duration (right panels) for the EVHOE 2003-2017. Upper panels are for the IBTS longer 100m sweeps configuration for EVHOE. The poorest survey coverage was in 2007, in contrast 2010 was one of EVHOE's best years for survey coverage.

While swept area is a common unit of standardization, incorporating multiple sources of observation error simultaneously is a newer and broader idea. An example approach was presented at IBTS in 2006 as part of an objective framework to confirm 'standard' (i.e. valid) hauls. Multiple trawl and oceanographic parameters such as trawl speed, opening, spread, angle as well as wave height, windspeed, warp and other data were analysed for six Swedish groundfish surveys. Catch data were also included later in the study. Data analysed was for hauls already identified as valid and used as part of the normal assessment process. Various clustering and multivariate techniques were applied, but even a simple PCA illustrated that for valid hauls, when multiple parameters were taken in combination, a significant number of hauls separate from the normal distribution.

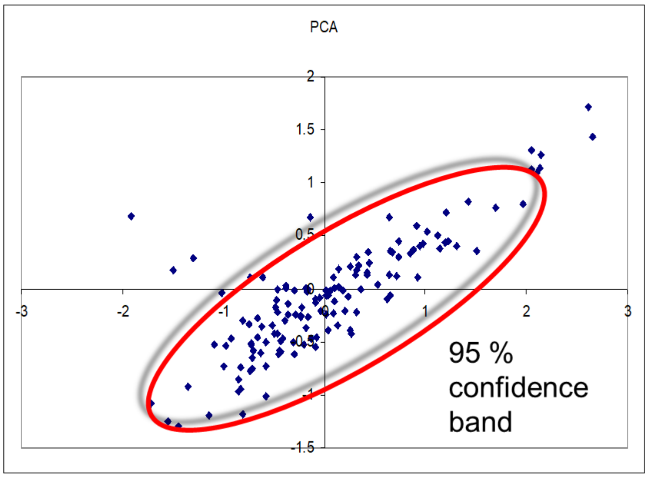


Figure 3.2.4 PCA of multiple trawl and oceanographic parameters for valid hauls from 6 Swedish groundfish surveys.

The extent to which interactions in potential 'outlier' hauls may affect the survey catchability is very difficult to quantify. However, monitoring complex physical sampling conditions with accuracy is ongoing in most surveys on an annual basis. Having a real time decision framework to evaluate how the survey is operating across multiple factors, even in a relative sense, compared to previous hauls and years was felt useful by the group. Decisions then can be made to either exclude certain points based on the output value for a given haul or include all data and provide the overall value for the survey as possible weighing or estimate of observation error.

Quantifying observed sampling conditions also takes pressure off survey scientists from regularly having to question at what point they should stop sampling, which is a common issue. Given the likely effect on survey observation error, a transparent and objective framework for quantifying sampling conditions beyond the control of the survey team seems to have merit.

4 Additional information to improve survey indices

4.1 Swept area

Discussion around reducing noise in survey sampling largely revolved around work ongoing at IBTS/DATRAS anyway in terms of standardizing survey data by swept area. This was expanded upon under Tor A (see Section 3.2) as regards other trawl and environmental covariates. If and how a broader range of variance information could be incorporated with survey data or into assessments was felt worthy of further consideration once we knew the direction the assessments were taking post benchmark. It was agreed however, swept area would be a useful progression and achievable in preparation time available for the upcoming benchmark. Relevant data exchange for updating the indices was identified to be carried out over the coming weeks between IFREMER and MI.

4.2 Cod tagging studies

A tagging study of Irish Sea cod began in 2016 in part to address issues around stock structure and migration (see WKIrish2[5]). The key issue being the recent, but ongoing disappearance of cod age 3+ from the Irish Sea 27.7a assessment. As a result, the assessment model has been strongly revising these year classes, and SSB, downward causing a significant retrospective bias problem.

Up to January 2019 4,238 cod were caught and tagged aboard chartered commercial fishing vessels using semi-pelagic fishing gear as part of a fisheries science partnership (FSP) survey. In addition, fish for tagging were sourced from shore angling competitions and others. Up to January 2019 138 tagged cod were returned (Fig 4.2.1).

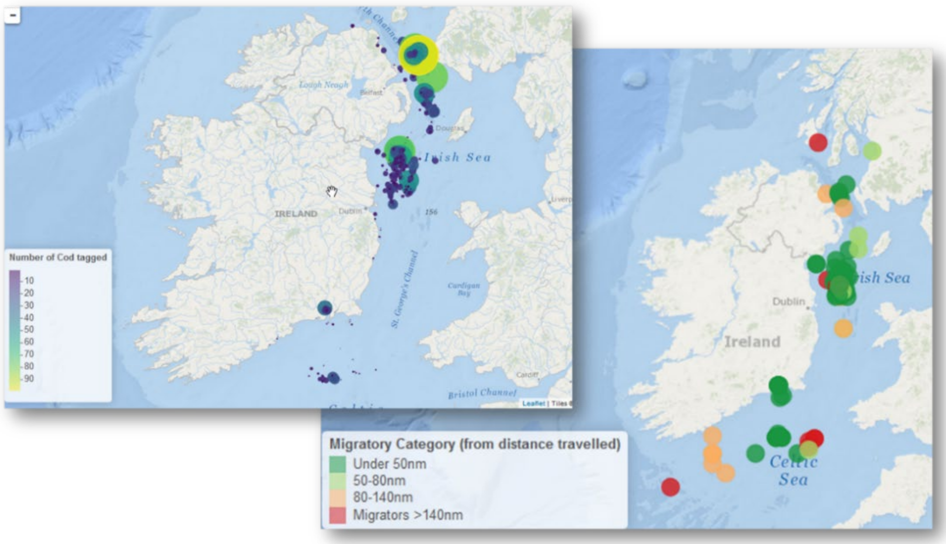


Figure 4.2.1 Release (left panel) and recapture sites (right panel) for Irish Sea cod tagging project.

The project relies on collaboration with the fishing industry to provide the data to develop a better understanding of the behaviour, biology and stock status of Irish Sea cod. Most recent

[5] ICES. 2016. Report of the Data Evaluation Workshop on Irish Sea Fisheries (WKIrish2), 26–29 September 2016, Belfast, UK, ICES CM 2016/BSG:02.

results suggest a strong migratory behaviour of Irish Sea cod into the Celtic Sea, suggesting up to 18% of mature fish migrate out of the Irish Sea.

Spawning Stock Biomass (SSB) and Fishing mortality (F) were evaluated under two different migratory hypotheses and compared with the model used at WGCSE 2018[6]. Migration was adjusted for the three most recent years (2015-2017) and give quite different perceptions of the stock (Fig 4.2.2).

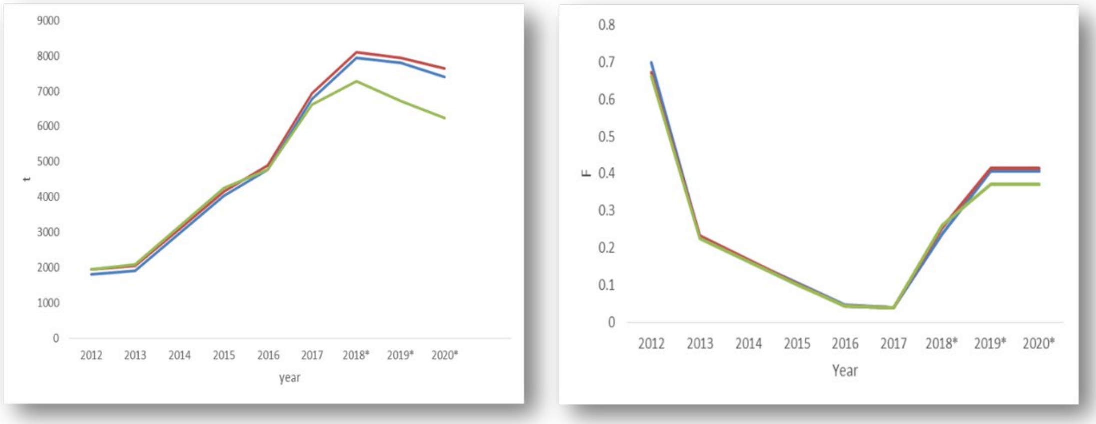


Fig 4.2.2 Hypothesis 1 - Blue: Migratory stock that returns to spawning sites in 7a (taking into account catches of 4+ year old fish in are 7g); Hypothesis 2 - Green: Emigration out of 7a (adjusted M); Red - WGCSE baseline (stock assessment without considering migration).

This is important information for survey design and indeed the future management and assessment of both stocks, but additional research is necessary.

Further work is essential to better understanding stock structure and improve future management. This includes investigation of migration as well as natural mortality in the Irish Sea. It might be necessary for a combined approach to manage the stocks in 7A and 7E-G, in light of the recent tagging study results.

[6] ICES. 2018. Report of the Working Group on Celtic Seas Ecoregion (WGCSE), 9–18 May 2018, Copenhagen, Denmark. ICES CM 2018/ACOM:13. 1887 pp.

5 Modelling approaches to:

5.1 Evaluating survey performance at the assessment level

Cefas has been working towards a framework for evaluating the impact of survey design changes on the advisory processes. Historically, the focus of survey data collection has been on consistency through time. However, invariably surveys will be faced at some point with managing the impact of wanted or unwanted changes that arise. Be that vessel breakdown, extreme weather or the progression to a new vessel or sampling equipment etc.. Resource pressures and changing management strategies may also require realignment of survey objectives and design. The aim of the project is to address impacts at the level of the management metrics, i.e. F, SSB and recruitment coming out of the assessment process. It replicates the actual management advice process from data collection, including biological sampling through assessment.

The project is still in its early development and so far efforts have focused on three aspects, first implementing different ways of subsampling, second implementing different methods of index calculation, both model and design based indices and third automating the assessment process in an effective way to allow the evaluation of large numbers of simulations. As an example of what might be done, preliminary analysis comparing two methods of survey effort reduction was presented to evaluate the impact of reducing tow duration from 30 minutes to 15 minutes. One method halved tow duration and the second method halved number of tows.

Methods of index calculation were i) current ICES/DATRAS index used for haddock and whiting in so far as could be replicated; ii) the delta-gam method was used for cod. For simulations using only 50% of the stations, stations were randomly subsampled from a post stratification scheme based on ecological strata.

Preliminary results suggest:

1. neither method of effort reduction had a major effect on the variability of assessment outputs, but unexpectedly some biases developed compared to the current assessment based in both cases.
2. ICES indices could not be replicated due to a subjective choice of data to include in specific years (mainly in the historic part of the series) so that effects of sampling at the station level required implementation of consistent ALK aggregation so that the overall index differed from that officially provided by ICES to the assessment process.
3. the consistency of the results when analysed with the surba method in parallel imply that the divergence is due to a shift in the survey weighting within the assessment due to a reduction in the abundance of older individuals in the survey indices.
4. the impact of using different methods for survey index calculation had a bigger effect on the assessment outcomes than did the reduction in sampling effort.

Given the simulations for the current assessment methods (Figure 5.1.1), one would conclude that from an efficiency perspective a 50% reduction in the number of stations is preferable to a 50% reduction in tow duration (Figure 5.1.2), particularly since from a practical perspective reducing the number of stations saves significantly more ships time than the reduction of tow duration.

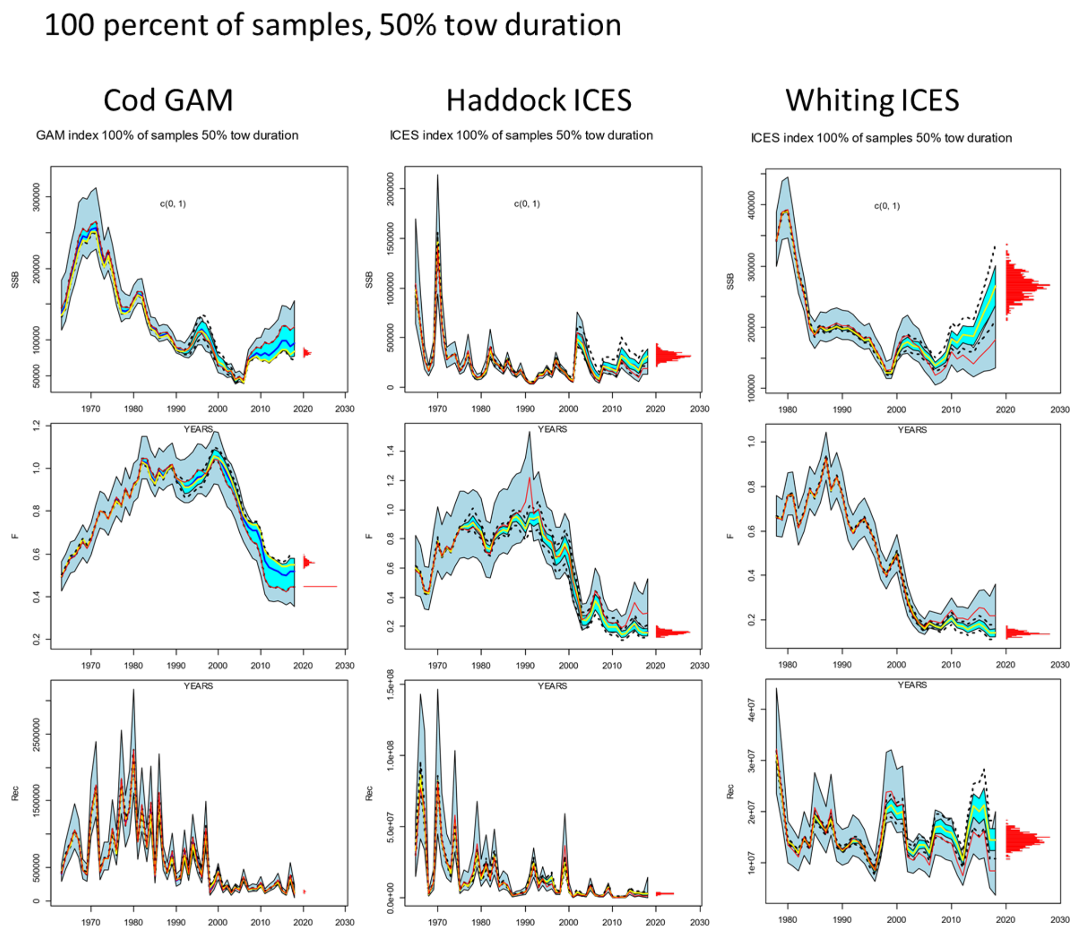


Figure 5.1.1: Stock assessment output for three species (NS cod, haddock, whiting left to right, SSB, F, recruitment top to bottom) compared to the current stock assessment as performed by WGNSSK 2018 using the approved index method for each assessment for 50% of stations at full tow duration.

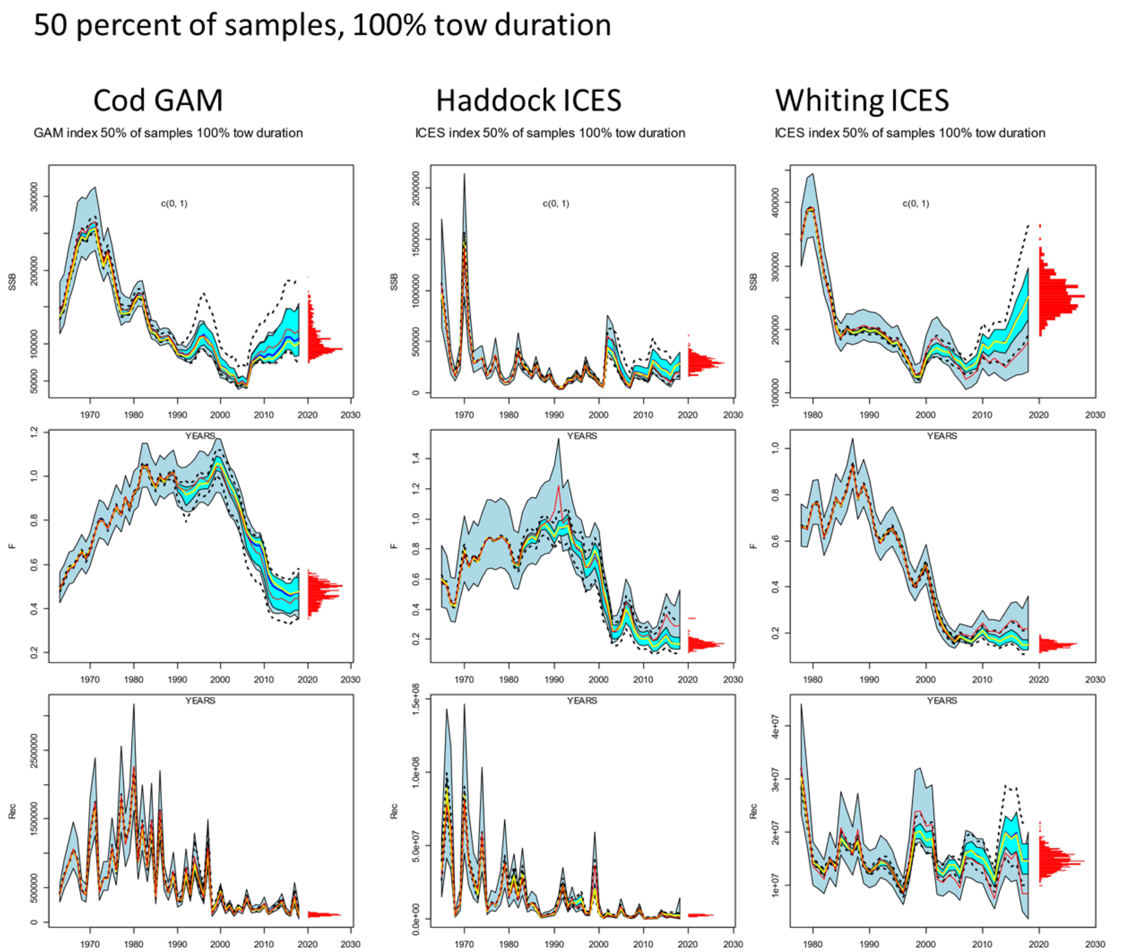


Figure 5.1.2: Stock assessment output for three species (NS cod, haddock, whiting left to right, SSB, F, recruitment top to bottom) compared to the current stock assessment as performed by WGNSSK 2018 using the approved index method for each assessment for all of stations at 50% tow duration.

Red line and grey-blue polygons indicate the estimate and the uncertainty according to the current stock assessment. Yellow and blue lines show the median and mean of the 500 simulations, with cyan polygons and dashed lines indicating 95% ci and min and max. Red histograms indicate the distribution of the 500 results.

5.2 Uncertainty estimation around age sampling

The work presented is based on North Sea IBTS data, but applicable to NE Atlantic stratified random estimates also. Three areas of project focus are:

1. The development of spatial age–length key (ALK) estimators for estimating indices of abundance at age
2. The development of uncertainty estimators for abundance at age indices and ALK estimators
3. Investigating the effect on abundance at age and its uncertainty if fewer otoliths are sampled or the number of hauls is reduced or increased

The survey design of the North Sea IBTS is based on stratification at two levels: roundfish areas (RFAs) and statistical rectangles (Figure 5.2.1 (a)). However, ICES DATRAS uses a post-stratification of index areas for indices calculation for each target species. These are based on aggregated RFAs rather than potential drivers of spatial distribution and variability such as bathymetry Figure 5.2.1 (b).

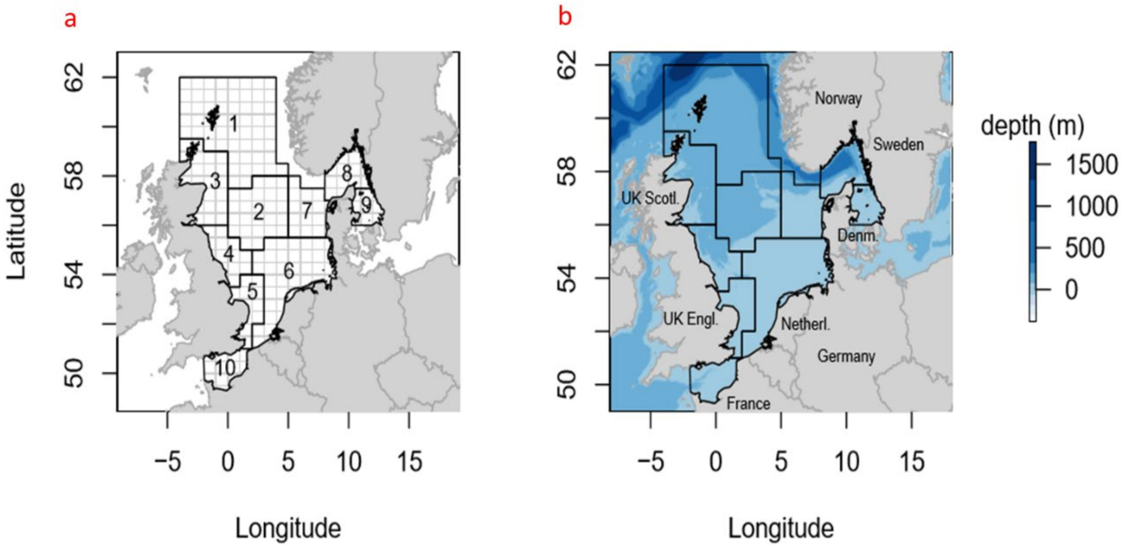


Figure 5.2.1: Standard roundfish areas (RFAs) used for roundfish since 1980 and for all standard species since 1991 (a - left panel). The number 1, for example, indicates ICES RFA 1. The small grey rectangles indicate the statistical rectangles of approximately 30 x 30 nautical miles (these vary from approximately 28 nm wide in the north, to approximately 40 nm wide in the south of North Sea). Right panel (b) shows underlying seabed topography including Norwegian Trench and shelf edge.

ICES DATRAS assumes the age–length relationship is constant over large areas. Violations of assumption will give biased results. Using IBTS cod data for Q1 in 2018 we can see quite pronounced spatial patterns to length-at-age even within RFAs (Fig 5.2.2).

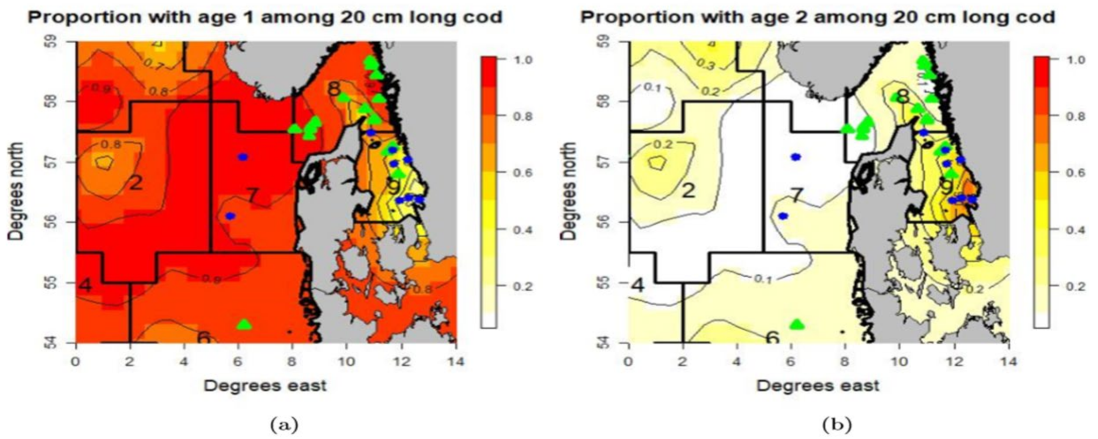


Figure 5.2.2: Estimated probability of age 1- and 2-year-old cod at length 20-cm cod in the first quarter of year 2018. Green triangles are observations of 1yr old fish; blue circles are observations of 2yr old fish in this length class.

DATRAS Area based ALK

An ALK is calculated for each RFA, and the index is calculated by taking the mean catch-at-age per rectangle, and then the mean over all rectangles in the index area. We refer this ALK estimator as the **Area based ALK**:

- A **single** ALK is produced for a RFA
- "borrowing" of ages from length classes elsewhere in RFA
- Missing ALK data are imputed by extrapolation: taking averages and borrowing from closest length group
- Variance estimated by bootstrap with replacement of randomly selected hauls within a RFA[7].

Issues:

- biased estimates (Berg et al., 2012[8], 2014[9]; Gerritsen et al., 2006[10]; Kimura 1997[11])
- Ignores spatial variation

We have proposed two approaches for maintaining the spatial integrity of the data. First, a spatial ALK estimator, where an ALK is produced for each haul and spatial variation in the data therefore maintained in the model. This we refer to as a **Haul based ALK**.

Haul based ALK

- borrow from same haul with length of $\pm 1\ cm$; if not possible
- borrow from closest haul (less than 60 nautical miles away); if not possible
- Find an alternative way to account for missing data (Model based ALK).
- Variance is estimated using nonparametric *stratified* version of bootstrapping approach taken by DATRAS

Model based ALK

A Multinomial logistic model is used to predict the distribution of proportion at age. Using a Gaussian Random Field to obtain smooth structures in the age–length distribution and linear predictor assuming a zero mean Gaussian spatial random field (models spatial effect).

$$p_a(l, s) = \frac{\exp[\mu_a(s,\ l)]}{1 + \sum_{i=M+1}^{A-1} \exp[\mu_i(l,\ s)]} \qquad \text{for } a = M+1,\ \ldots,\ A-1$$

$$\mu_a\,(s,\ l) = \ f_a(l) + \gamma_a(s)$$

- $\boldsymbol{p_a(l, s)}$: probability of a fish with length l at location s to be of age a

---

[7] Report of the Workshop on Implementation in DATRAS of Confidence Limits Esti-mation of Abundance Indices from Bottom Trawl Survey Survey Data. *International Council for the exploration of the Sea*, ICES DATRAS REPORT.

[8] http://dx.doi.org/10.1016/j.fishres.2012.06.016

[9] http://dx.doi.org/10.1016/j.fishres.2013.10.005

[10] doi:10.1016/j.icesjms.2006.04.008

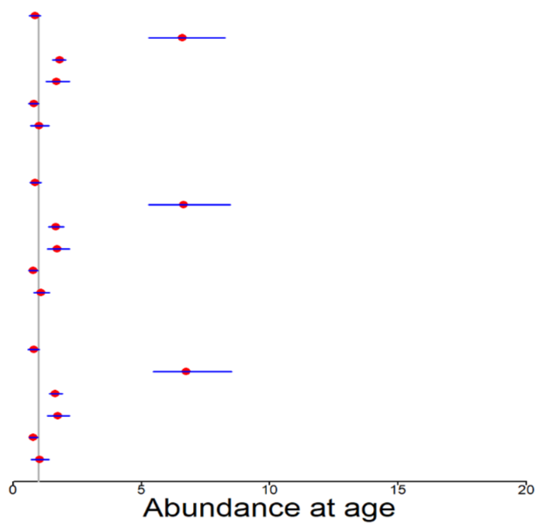
[11] DOI: 10.1080/10641260600621761

- Either:
  - $p_A(l, s) = 0$ or
  - $p_M(l, s) = 0$, and the other is selected such that $\sum_a p_a = 1$
- $\boldsymbol{f_a(l)}$: is the a continuous function of length l
  - Is intended to account for the fact that longer fish are typically older, and is modelled with P-splines
  - P-splines are included as a zero mean Gaussian random effect
- $\boldsymbol{\gamma}$: is a zero mean Gaussian spatial random field with Matern covariance function
  - is intended to account for the spatial variation in the ALK
- Variance is estimated by stratified bootstrap procedure, where hauls in a rectangle are sampled ran-domly with replacement and ages in the resampled hauls are taken.

Applying 150 bootstraps and comparing estimates across the three ALK methods shows some small changes in age structure, but nothing too worrying. The most noticeable contrast is the reduced variance with the DATRAS Area based method as it lacks the ability to account for spa-tial variation in the data.

Table 5.2.1 Estimates of mean CPUE at age (with 95% confidence intervals) for cod in the North Sea using 3 alternative ALK derivation methods.

Age	Species	SE	RSE%	Width CI-95	Estimate (CI_{95})
Area based					
1	cod	0.124	14.49	0.428	0.86 (0.65-1.07)
2	cod	0.872	13.21	2.959	6.60 (5.29-8.25)
3	cod	0.174	9.62	0.491	1.81 (1.55-2.04)
4	cod	0.281	16.50	0.890	1.70 (1.30-2.19)
5	cod	0.124	15.39	0.371	0.81 (0.62-0.99)
6+	cod	0.244	23.81	0.695	1.02 (0.69-1.39)
Haul based					
1	cod	0.126	14.57	0.437	0.87 (0.65-1.09)
2	cod	0.936	14.07	3.173	6.65 (5.28-8.46)
3	cod	0.162	9.68	0.587	1.67 (1.39-1.97)
4	cod	0.283	16.42	0.855	1.72 (1.34-2.20)
5	cod	0.117	14.88	0.337	0.79 (0.61-0.95)
6+	cod	0.267	24.30	0.610	1.10 (0.81-1.42)
Model based					
1	cod	0.124	15.21	0.433	0.82 (0.59-1.02)
2	cod	0.918	13.60	3.042	6.75 (5.46-8.50)
3	cod	0.178	10.77	0.501	1.65 (1.42-1.92)
4	cod	0.292	16.83	0.853	1.74 (1.35-2.20)
5	cod	0.115	14.59	0.319	0.79 (0.63-0.95)
6+	cod	0.252	24.01	0.675	1.05 (0.71-1.39)



As a follow up to the above discussion, the modelled ALK method was applied to IGFS haddock data for a number of hauls. This was to compare with the similar approach[12] currently used for filling in ALKs for the cod, haddock and whiting combined indices. Modelled ALKs for two of the example hauls are given in Fig 5.2.3., as well as a map of the probability of a 30cm fish being age 2 over the full survey area. Fitting at the haul level was good, although some reduction in

[12] ICES Journal of Marine Science, 63: 1096e1100 (2006) doi:10.1016/j.icesjms.2006.04.008

size from Age 4-5 was apparent, possibly due to reduced data at those ages. The overall survey model is under development and it was discussed that the Stochastic Partial Differentiation Equation (SPDE) process within the model easily facilitates different ranges on landvs.at sea and this would be a good next step to improve these overall survey level fits.

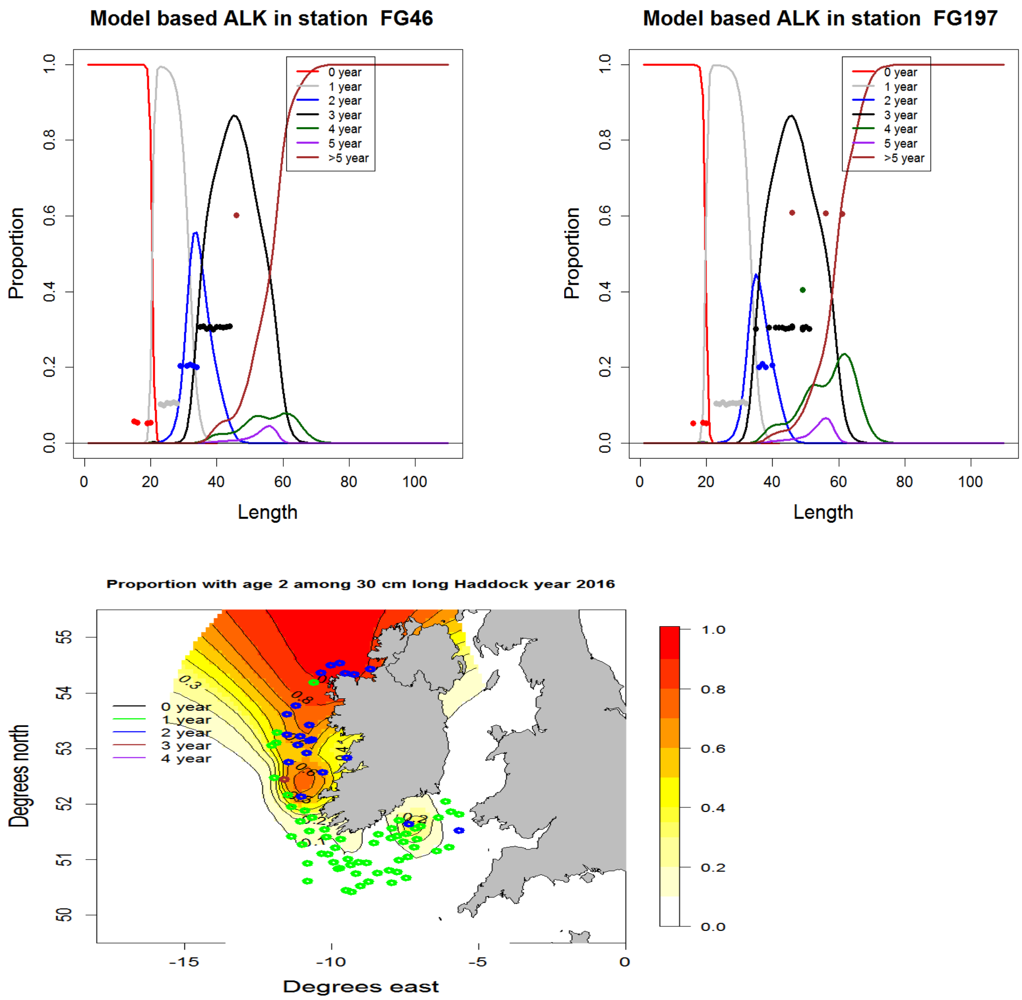


Fig 5.2.3: Top panels show modelled ALK probability (solid lines) from observed age at length (solid dots) for IGFS stations FG46 and FG197 data in 2016. Lower panel shows the probability of a 30cm haddock being Age 2 over the full survey area. Probabilities are shown as contours; observations at age are solid dots.

5.3 Minimizing the impact of missing survey data

A number of motivations lead to considering modelling as an alternative method for index calculation. The main ones being current index derivation methods for Celtic Sea cod, haddock and whiting do not:

i. Compensate for missing survey hauls due to vessel breakdown, poor weather etc.
ii. Currently output estimates of uncertainty
iii. Afford inclusion additional information on catchability or density covariates of interest (not covered here)

While point ii) is not strictly a modelling issue, the need for work in this area was highlighted in 2017 when the EVHOE had significant vessel problems and survey coverage in the Celtic Sea for that year was greatly reduced (Fig 5.3.1). This, combined with a particularly poor short term management forecast for Celtic Sea cod (WGCSE2018[13]) naturally focused attention on all assessment inputs.

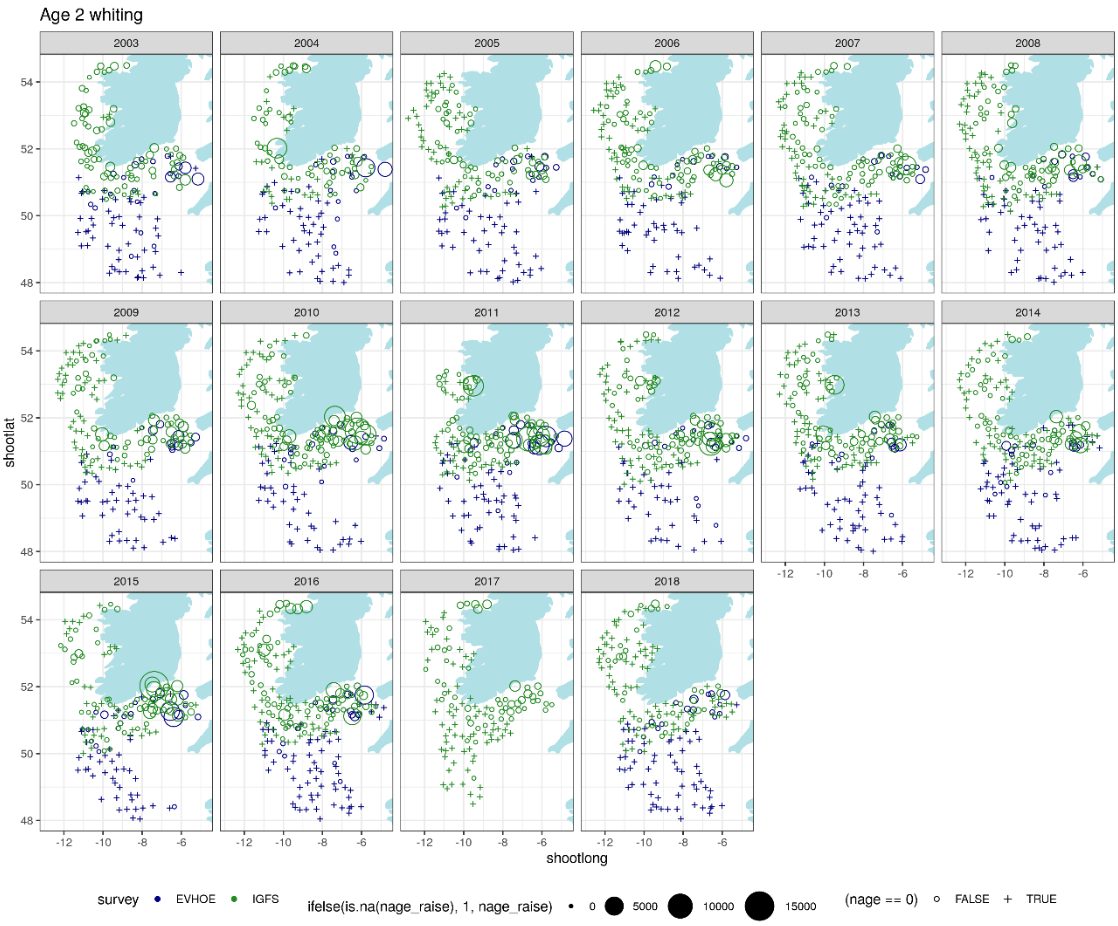


Fig 5.3.1 Maps of station positions for each survey with null catches of Age 2 whiting indicated by a cross, positive EVHOE catches = blue circle, IGFS catches = green circle. Reduced coverage in 2017 clearly seen as well as some more southerly hauls added by IGFS to extend coverage.

Geostatistical methods can offer some advantages over conventional design-based approaches as they predict population densities at nearby sites to be more similar than densities at geographically distant sites. Hence, the deviation of observed catches from the expected (i.e. model prediction) at a location is hopefully reduced, leading to overall increased precision[14] for the observed and interpolated areas. The ability add covariates to further explain spatial patterns in catch abundance can enhance model fitting and precision of the index further again. In contrast, stratified design based estimators assume density is constant within the given strata and variance estimated as deviations from the overall strata mean.

[13] https://doi.org/10.17895/ices.pub.4490

[14] Thorson, James T., Andrew O. Shelton, Eric J. Ward, and Hans J. Skaug. 2015. 'Geostatistical Delta-Generalized Linear Mixed Models Improve Precision for Estimated Abundance Indices for West Coast Groundfishes'. *ICES Journal of Marine Science* 72(5):1297–1310. doi:10.1093/icesjms/fsu243

The work presented is for the VAST (Vector Autoregressive Spatio-Temporal) model (www.github.com/james-thorson/VAST). VAST is a spatially explicit model that predicts population density for all locations over the spatial domain. It derives quantities (e.g. biomass, abundance) by aggregating this predicted density across the spatial domain, weighted by the area associated with each estimate. VAST also interpolates biomass for unsampled areas using spatially correlated random effects.

Data

Survey data comprise numbers at length $Y_{t,s,h,l}$ (indexed by: year t, survey s, haul h, length class l). Numbers at length are converted to numbers-at-age using survey-specific age–length keys

$$k_{t,s,j,l,a} : y_{t,s,h(j),l} \rightarrow \check{y}_{t,s,h,a}$$

where $k_{t,s,j,l,a}$ is the ALK giving the probability *a* fish of length *l* is age *a* in a given survey and Stratum j. Note, to-date we do not include stratum-specific EVHOE survey ALKs. Response used in the geostatistical models below is thus the number-at-age per survey and haul denoted $\check{y}_{t,s,h,a}$. The data used here is for Celtic Sea whiting.

Model structure

The Vector Autoregressive Spatio-Temporal (VAST) model is a flexible tool for modelling univariate and multivariate spatio-temporal species distributions[15]. VAST implements a delta-model capable of dealing with zeros and a continuous positive distribution. We fit VAST separately for each age and examples here are for Age 2. The two components of the model fit to each age are:

- Encounter probability (Bernoulli)

  $$P(\check{y}_{t,s,h,a} > 0) = P_{t,s,h,a}(\theta_p)$$

- Positive catches

  $$P(\check{y}_{t,s,h,a} \mid \check{y}_{t,s,h,a} > 0) = f(\check{y}_{t,s,h,a} \mid \theta_f)$$

  where θ_f is a set of parameters governing the probability distribution function f.

Encounter rate linear predictors incorporate year-effects, survey differences, spatial and spatio-temporal variability. Positive catches predictions are structured similarly other than a gamma distribution assumption, log-link function and swept area offset.

We defined the spatial domain as two hundred knots implemented using k-means clustering to give knot positions proportional to sampling intensity (Thorson, 2019). VAST is estimated using the Laplace approximation to the marginal likelihood in TMB[16] with the conditional probability

[15] Thorson, James T. 2019. 'Guidance for Decisions Using the Vector Autoregressive Spatio-Temporal (VAST) Package in Stock, Ecosystem, Habitat and Climate Assessments'. *Fisheries Research* 210:143–61. https://doi.org/10.1016/j.fishres.2018.10.013

[16] Kristensen, Kasper, Anders Nielsen, Casper W. Berg, Hans Skaug, and Brad Bell. 2016. 'TMB: Automatic Differentiation and Laplace Approximation'. *Journal of Statistical Software* 70(5).

of random effects approximated using the stochastic partial differential equation (SPDE in R-INLA[17]).

Output summary

The predicted encounter probability and positive catch rates show reasonably good correlation with the observations for most age classes with Age 2 shown in Fig 5.3.1. Log density of abundance (Fig 5.3.2) shows spatial and temporal patterns similar to the observed catches (presented above) as well as evidence of the strong 2009 and 2011 years classes (plus 2-year lag).

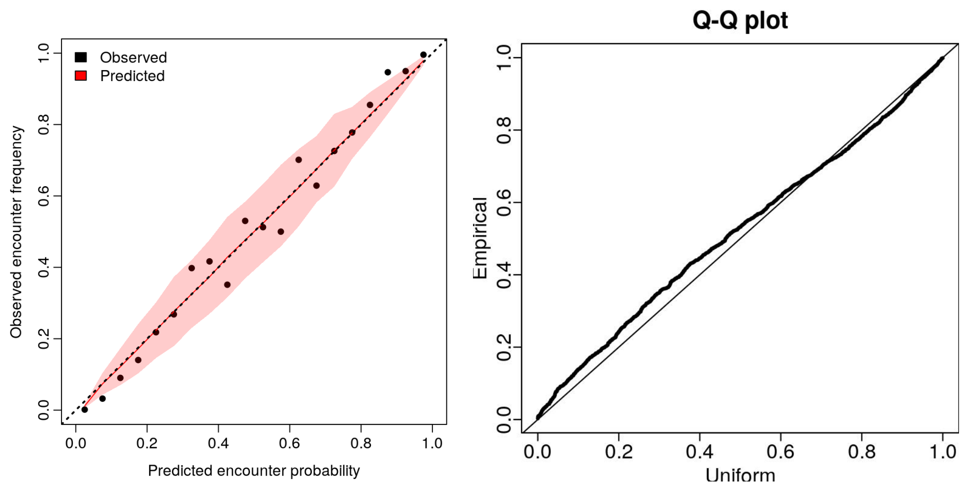


Fig 5.3.1 Whiting in area 7b-ce-k age 2: residual diagnostics showing predicted encounter probability against observed (left panel); QQ plot for positive catches (right panel).

[17] Rue, Håvard. n.d. 'Bayesian Spatial Modelling with R-INLA'.

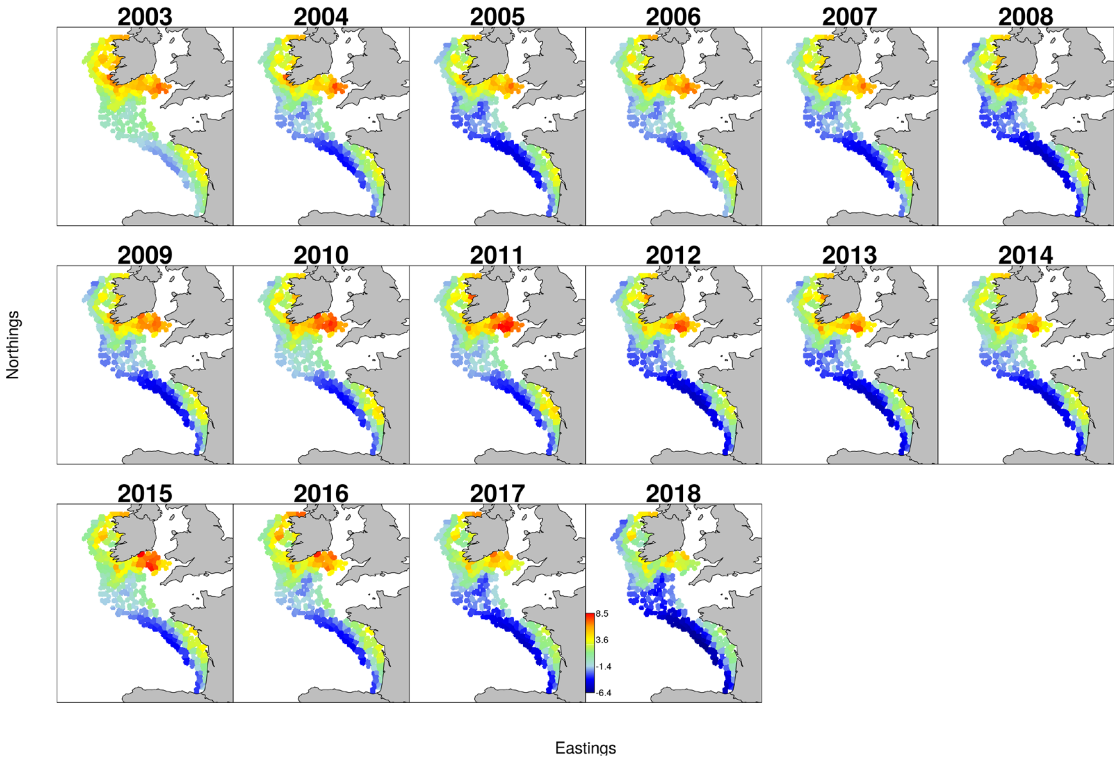


Fig 5.3.2 Whiting in area 7b-ce-k age 2: Predicted log density over space and time. Strong recruitment in 2009 and 2013 can be seen as 2 yr old hot-spots in 2011 and 2015 respectively.

Transition of strong cohorts can be seen to transition clearly through the index (Fig 5.3.3). The first important sanity check was to compare the VAST and current indices by age directly (Fig 5.3.4) which showed good correlation again.

A further important check of model stability we felt was to sequentially remove the EVHOE survey data to mimic what had been lost in 2017. This ensured we could test model stability with loss of a full survey over a range of conditions; especially high and low catch years, against the original index using the full set of observed data (Fig 5.3.5). This was to confirm that the model, as configured, was able to use the spatial and temporal information available as well as vessel/catchability effects to '*fill-in*' reasonable catch density estimates.

The VAST model provided very similar predictions for the Age 2 whiting relative to the original observed data overall (Fig 5.3.5). Where we start to see some deviation is where prediction for large data gaps is required in a year of unusually high (and presumably low) survey index. Intuitively then, the more the neighbouring years and areas of a survey resemble the year of missing data, the more accurate the model predictions are going to be. It is important therefore, in bringing an index forward into the assessment process, that consideration be given to the uncertainty estimate as well as the index itself. In that respect VAST does explicitly model habitat and catchability covariates separately so that vessel effects etc. are "controlled for" and density predictions are influenced by habitat drivers only.

Discussion of the VAST method and model specification centred on checking error distribution particularly in relation to over dispersion to ensure central tendency not being dragged up. One suggestion was to look at larger polygons to include more data points within each polygon. Spatial random effects could also be included to tune model fit using depth and environmental covariates. In addition look further into how VAST is or can deal with isotropy.

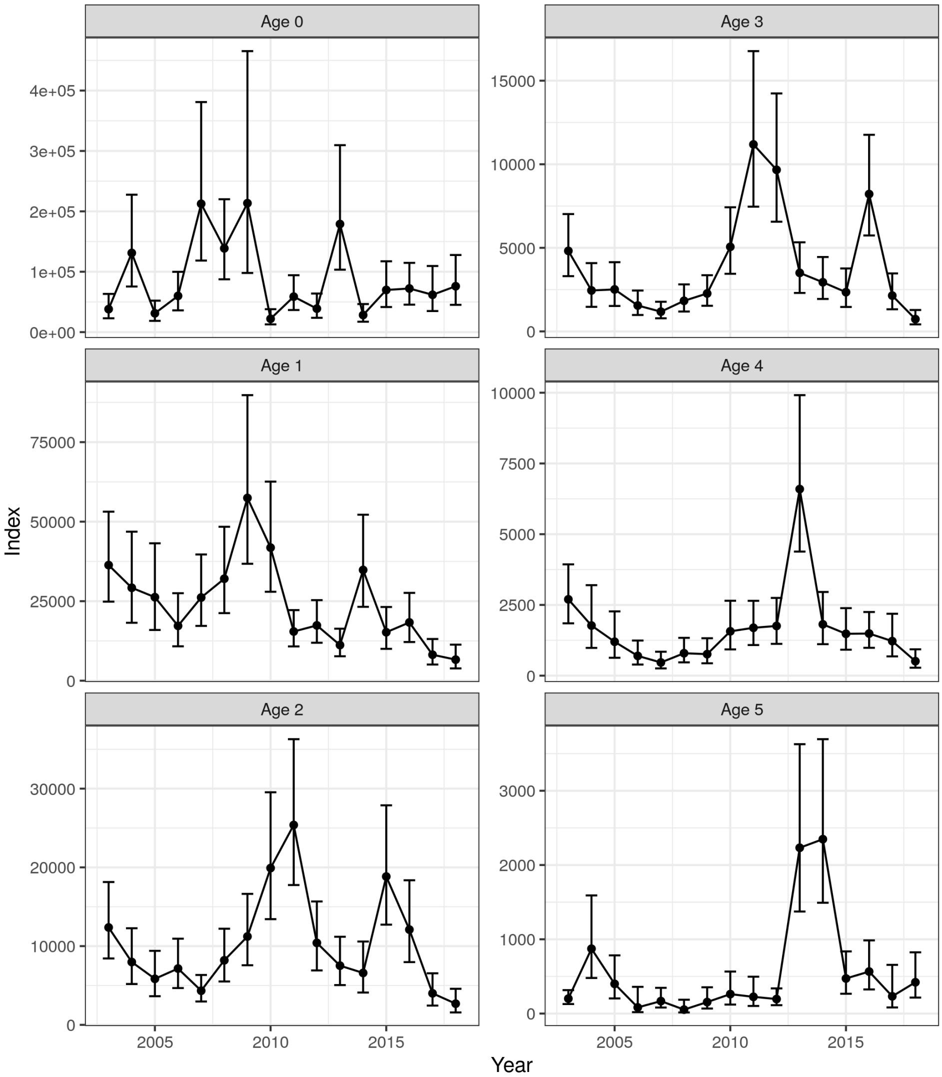


Fig 5.3.3 Whiting in area 7b-ce-k: Estimated index and uncertainty for all ages included in the assessment (0-5+).

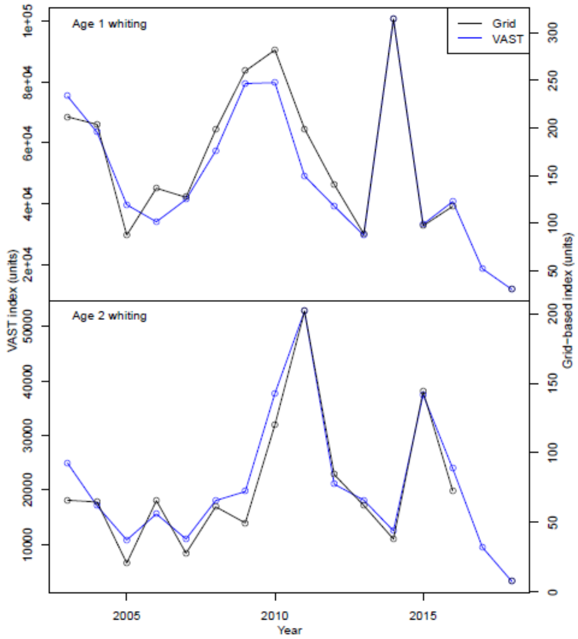


Fig 5.3.4 Whiting in area 7b-ce-k age 1 - 2: Relative comparison of VAST model predicted index vs. current grid approach for age 1 (top panel) and age 2 (bottom panel).

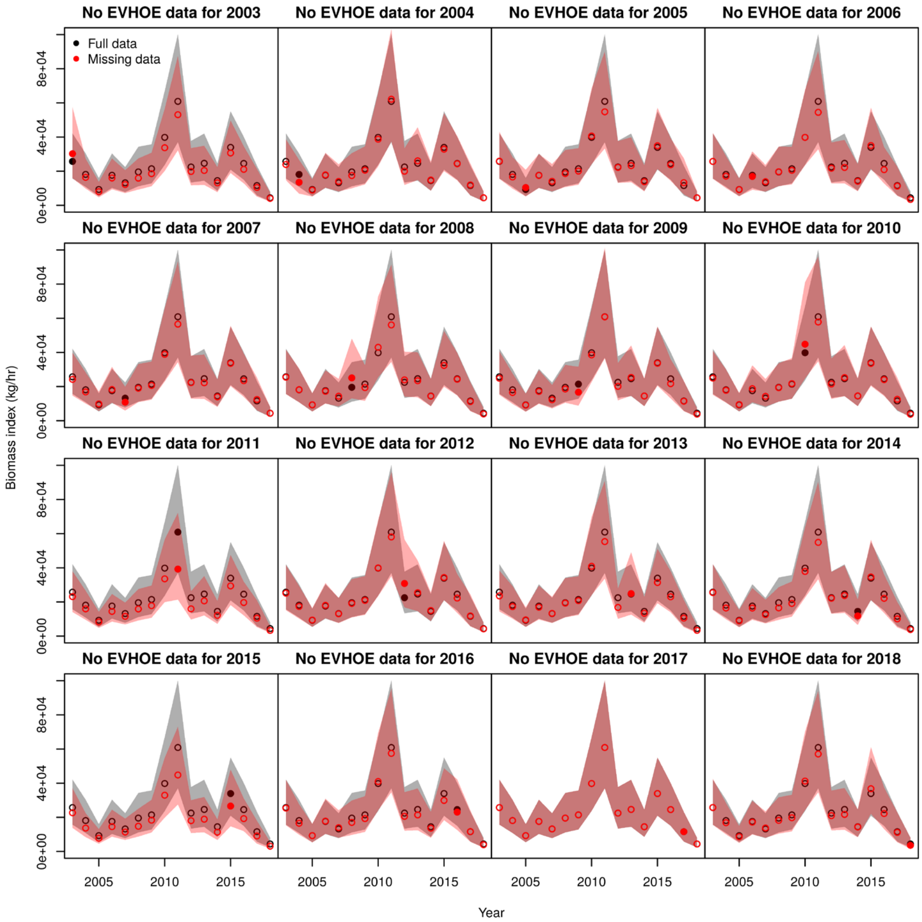


Fig 5.3.5 Whiting in area 7b-ce-k age 2: The impact of missing data were investigated by sequentially removing the EVHOE survey data year on year to mimic survey failure for that year and impact on the time-series. The index based full dataset

is given as black circles with grey confidence intervals. The recalculated index with a missing year is given as red empty circles with the red solid circle indicating the annual data point where EVHOE survey has been removed. Confidnce bounds are shaded light red for this recalculated index. The VAST model produced very similar biomass estimates for the full dataset and the partial data, suggesting that the survey coverage in 2017 was sufficient for the model to accurately estimate the index. It is noteworthy that when data is removed from a year with an unusually high index that the re-estimated index shows the greatest difference from the original.

Annex 1: List of participants

Name	Institute	Country (of institute)	Email
Olav Nikolai Breivik	IMR	Bergen, Norway	Olav.Nikolai.Breivik@nr.no
Paul Dolder	CEFAS	Lowestoft, UK	paul.dolder@cefas.co.uk
Hans Gerritsen	The Marine Institute	Galway, Ireland	hans.gerritsen@marine.ie
Natoya Jourdain	IMR	Bergen, Norway	natoya.jourdain@hi.no
Cóilín Minto	GMIT	Galway, Ireland	Coilin.Minto@gmit.ie
Lionel Pawlowski	IFREMER	Lorient, France	Lionel.Pawlowski@ifremer.fr
Sven Kupchus	CEFAS	Lowestoft, UK	sven@kupschus.net
Marianne Robert	IFREMER	Lorient, France	Marianne.Robert@ifremer.fr
David Stokes	The Marine Institute	Galway, Ireland	david.stokes@marine.ie

Annex 2: Resolutions

WKESIG – Workshop on evaluating survey information Celtic Sea gadoids

2017/2/EOSG08 The **Workshop on evaluating survey information Celtic Sea gadoids (WKESIG)**, chaired by David Stokes, Ireland, will meet in 4-5 February 2019 at the at the Marine Institute, Galway, Ireland to:

d) Review and consider the quality and availability of survey data going into the assessment of cod, haddock and whiting as requested by WGCSE2017 (Science plan codes 5.1);

e) Evaluate the potential to improve current survey indices by use of additional information such as standardizing by swept area or using model based index approaches (Science plan codes 5.1);

f) Review and standardize methods for evaluating and constructing indices including applying and filling in ALKs, estimating uncertainty and, where desirable, combining or complementing surveys with other data sources (Science plan codes 5.1).

WKESIG will report by 19 March 2019 for the attention of ACOM and SCICOM.

Supporting information

Priority	The Benchmark process is critical to the review and quality assurance of stock assessments within ICES. A number of points for investigation have been indentified by WGCSE for cod, haddock and whiting which form a key mixied fishery in the Celtic Sea. These stocks are largely, if not exclusively tuned using survey data and therefore this work is considered a high priority.
Scientific justification	Term of Reference a) Three of the largest stocks assessed by WGCSE (Cod 7e_k, Haddock 7b_k and Whiting 7b_k) form part of a significant mixed demersal fishery in the Celtic Sea. These assessments rely heavily on survey indices which in all cases use at least one combined survey index between Ireland (IE-IGFS) and France (EVHOE). How these data are standardized and combined is somewhat different across stocks and achieved by R code passed through earlier Benchmarks, but not currently published or documented in detail. The proposed workshop will review the construction and quality of these survey indices including: a) optimal standardization such as swept area b) appropriate fill ins of ALKs c) appropriate estimates of uncertainty which could be passed in to the assessment process.
Resource requirements	The input data for this work is largely available already through DATRAS and Intercatch and therefore time to process the inputs is main resource requirement.
Participants	The Group is normally attended by 6-10 members and guests.
Secretariat facilities	None.
Financial	No financial implications.
Linkages to advisory committees	The results and conclusions will likely feed into future benchmark processes for these or other species.
Linkages to other committees or groups	There is a close working relationship with IBTS, WGCSE and WGMIXFISH.
Linkages to other organizations	